8-2013

Targeted Deletion Of Functionally Validated Enhancers Defines Their Role in Mouse Limb Development

Mark J. Nolte

TARGETED DELETION OF FUNCTIONALLY VALIDATED ENHANCERS DEFINES THEIR ROLE IN MOUSE LIMB DEVELOPMENT

by

Mark Joseph Nolte, B.S.

APPROVED:

Richard R. Behringer, Ph.D.
Supervisory Professor

Michelle C. Barton, Ph.D.

Digitally signed by Yas Furuta
Date: 2013.08.09 17:29:49 +09'00'

Yasuhide Furuta, Ph.D.

James F. Martin, Ph.D., M.D.

Ambro van Hoof, Ph.D.

APPROVED:

Dean, The University of Texas
Graduate School of Biomedical Sciences at Houston

TARGETED DELETION OF FUNCTIONALLY VALIDATED ENHANCERS DEFINES THEIR ROLE IN MOUSE LIMB DEVELOPMENT

A

THESIS

Presented to the Faculty of The University of Texas Health Science Center at Houston and The University of Texas MD Anderson Cancer Center Graduate School of Biomedical Sciences in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Mark J. Nolte, B.S.

Houston, Texas

August, 2013

For Emily and Joyce

And for Greg

And for Suzanne and John

And for Simon, Allianna, Mary Tessa and Luna

"A long pull, a strong pull, and a pull all together."

~ Joseph Smith;

quoted by Brigham Young,

Deseret News: Semi-Weekly, Apr. 20, 1867, p. 2.

Acknowledgments

I will undoubtedly leave out the names of individuals who have positively impacted my experience in Houston, Texas. I regret this; if your name is not here please know that I will still recall your name, your face and your influence throughout my life. Please note that my gratitude is expressed in no particular order. Thanks to those who entered the GSBS Program in Genes and Development with me in 2006: Chih-Chao Yang, Amanda Brock, Gabriel Gonzalez, Jason Ford, Aaron Raymond, Hussein Abbas, Chris Gully. Chih-Chao and Amanda in particular, thanks for being interested in my life and my family. Thanks to all other colleagues in the Genes and Development Program, especially Le Huang, Svitlana Kurinna, Aundrietta Duncan and William Muñoz. Research is impossible without recreation. Thanks to the Texas Medical Center volleyball group: Svitlana, Sarah, Brittany, Derek, William, Jenna, Le, Vincent, Cheng-Chiu, Chih-Chao, Ka Chun and others on the sand. Thanks to the LDS Church volleyball group: Myles and Rishelle Scott, Robert and Erin Ross, Brittany and Derek Bitner, Chris and Sarah Hawks, T.C. and Emily Ence, Brian Crane, Brent and Kelly Woodward, Mike Hansen, Nate Ridge, Kevin Keys, Matt Sjoblom, Scott Ellsworth and others. Of course a big thanks to Richard Behringer, my mentor, for putting up with me in the lab and in his office. Especially important to me were the meetings, trips and lab outings Richard arranged that have strengthened my connections to wonderful and successful scientists. Thanks to all members of the Behringer Lab that I interacted with early on: Chris, Gabriel, Allison, Cheng-Chiu, Ken, David, Shannon, Liesl, Ellen, Peter, Ross, Chuan-Wei, Hao, Grant, Naomi, Bob, Ying and Jenny. And thanks to the "later" Berhinger Lab members: Gabriel, Allison, Rachel, Shuo-Ting, Zer, Ying and Jenny. Special thanks to Rachel – for your knowledge, experience, listening ear and sincere words of encouragement. Special thanks to Jenny – I could not have completed my project without your help; I will miss seeing and talking with you on a daily basis. Special thanks to Ying – without

you my project would not have come together; I will miss these last few months where we worked often with each other. Allison, we loved being friends with your entire family, and celebrating not a few birthdays through the years. Abel, Kenny and Ka Chun, I loved having you in the lab too for a short time. Thanks to all of my various committee members: Michelle Barton, Yas Furuta, Jim Martin, Ambro van Hoof, Randy Johnson, Miles Wilkinson and Michael Kohn. Thanks to Kendra Allton, Sabrina Stratton and Lindsey Minter for help with the qPCR. Thanks to Jan and Chad in the GEMF Core. Thanks to Hank Adams – we didn't do much microscopy together but we did a lot of talking; thanks for loving the outdoors and animals for the sake of animals. Thanks to Delfin, Ben, Lito and Lonny – my project would have faltered without your tireless assistance. Thanks to Alex, Douglas, Eunice, Winston and others for making my work possible everyday. Thanks Cordelia, James and Salpy. Special thanks to Salpy and Berge – how grateful I am for your help and friendship. Thanks to Stan Tsang for removing the demons from my computer and more importantly for the frienship. Thanks to Elisabeth Lindheim, the first friend I made in the Genes and Development Program – your kindness extended to our family via emails before I even arrived in Houston. Thanks for your example as a professional in a high-stress environment and as a friend on a daily basis. I will miss the car rides to the program retreats; and in the future I'm sure I will recall that much of what was positive about my entire time in Houston was a result of your steady friendship and concern. Thanks to David Caprette across the way at Rice University – I loved teaching and talking with you, even if it was just briefly after classes. Your emails towards the end of my Ph.D. were kind and most helpful. Thanks to Flavio Zolessi, and to all the fellow students of the 2012 PASI/LASDB Short Course in Montevideo, Uruguay. Your wit, enthusiasm and smarts provide the renewed strength I needed to finish. Special thanks to Leonel Muñoz, Alberto Stolfi and Jennifer Fish for their advice and encouragement. Leonel, I am so glad we met – how was I to know that I'd have to travel to another continent to meet such a dear, wise friend. Our

conversations after long days full of classes were and are so important to me. Thanks for walking and talking instead of taking a cab. On our behalf, here's thanks to Stephen Jay Gould. Thanks to Sarah Bentley for making bus rides into and from work enjoyable; our conversations made Houston into a hometown, and your advice about writing my thesis helped more than you probably know. Special thanks to Grant Adamson and Robert Ross and their families – our conversations about the Ph.D. experience were certainly as important as the degree itself. Grant, the influence of your friendship, sincerity and razor sharp intellect will be felt daily – though perhaps in a cognitively blended, unconscious manner. Perhaps most importantly, thank you for being kind to my family and to others because of what you learn. Thanks to Jennifer Moore – I learned so much from your relentless inventiveness, optimism and love for others. Thanks to my seminary students for filling my life with diverse friendships and laughs. Special thanks to Jenny Lieberum and Michael Mangum for being happy each morning; because you wanted me to be there I wanted to be there. Thanks to my parents, Greg and Joyce Nolte. Thank you for visiting, talking and guiding me. I owe much, if not all, of what good I contribute to this world to Greg and Joyce Nolte. Joyce, our conversations through the years have provided structure to my thoughts and goals; an answered phone call is an answered prayer. I recognize that my parents' love and support are also proxy for the love and support of countless family members – some whom I've met and some who have passed on many, many years ago – without whom I would not have even become a scientist. On that note, thanks to Marine and Robert Nolte for taking me to Sea World when I was six-years-old. Thanks to John, Brad, Matt, Julianna and Bryan and their families. I am the product of your positive influences. Matt and Julianna, thanks for maintaining interest in my research and aspirations; the two of you made distant states seem to have shared borders. Thanks to John and Suzanne Taye for never wavering in their support for me. Suzanne, what a joy your visits to Houston were, especially for Luna's birth. While here you lifted the burden that inevitably attends the long haul of research. Thanks to Simon for

being my close friend and fellow scientist, and for being a son while he's at it. Thanks, Simon, for believing in and loving your Dad. Thanks to Allianna, Mary Tessa and Luna for enriching my life with the support and love that can only bubble forth from a circle of sisters/daughters. Lastly, I thank my wife, Emily, for making life in Houston a life worth remembering and cherishing. I am grateful that we accomplish goals together; I am grateful that her support has been rock solid. Most certainly I would not have obtained this degree without Emily's perennial love, support and sheer exertion.

TARGETED DELETION OF FUNCTIONALLY VALIDATED ENHANCERS DEFINES THEIR ROLE IN MOUSE LIMB DEVELOPMENT

Publication No._____

Mark Joseph Nolte, Ph.D.

Supervisory Professor: Richard R. Behringer, Ph.D.

Transcriptional enhancers are genomic DNA sequences that contain clustered transcription factor (TF) binding sites. When combinations of TFs bind to enhancer sequences they act together with basal transcriptional machinery to regulate the timing, location and quantity of gene transcription. The diversity in morphological and behavioral traits between and within species is at least partially a result of the differential expression of individual genes and gene networks during embryological and postnatal development. Thus elucidating the genetic mechanisms responsible for differential gene expression is essential to an understanding of development, evolution and disease. Numerous methods are in use to identify and characterize enhancers. Several high-throughput methods generate large datasets of enhancer sequences with putative roles in embryonic development. However, few enhancers have been deleted from the genome to determine their roles in the development of specific structures, such as the limb. Manipulation of enhancers at their endogenous loci, such as the deletion of such elements, is essential to understanding the regulatory interactions, rules and complexities that contribute to faithful and variant gene transcription – the molecular genetic substrate of evolution and disease. To understand the endogenous roles of two distinct enhancers known to be active in the mouse embryo limb bud we deleted them from the mouse genome. I hypothesized that deletion of these enhancers would lead to aberrant limb development.

The enhancers were selected because of their association with p300, a protein associated

with active transcription, and because the human enhancer sequences drive distinct *lacZ* expression patterns in limb buds of embryonic day (E) 11.5 transgenic mice. To confirm that the orthologous mouse enhancers, mouse *280* and *1442* (*M280* and *M1442*, respectively), regulate expression in the developing limb we generated stable transgenic lines, and examined *lacZ* expression. In *M280-lacZ* mice, expression was detected in E11.5 fore- and hindlimbs in a region that corresponds to digits II-IV. *M1442-lacZ* mice exhibited *lacZ* expression in posterior and anterior margins of the fore- and hindlimbs that overlapped with digits I and V and several wrist bones. We generated mice lacking the *M280* and *M1442* enhancers by gene targeting, replacing the enhancers with a *neomycin resistance* gene that was subsequently removed by Cre recombinase. Intercrosses between *M280* -/+ and *M1442* -/+, respectively, generated *M280* and *M1442* null mice, which are born at expected Mendelian ratios and manifest no gross limb malformations. Quantitative real-time PCR of mutant E11.5 limb buds indicated that significant changes in transcriptional output of enhancer-proximal genes accompanied the deletion of both *M280* and *M1442*. However, when alizarin red and alcian blue stains were used to visualize bone and cartilage in neonatal null mice we observed that all limb bones were present in their expected positions, an observation also confirmed by histology of E18.5 distal limbs. Fine-scale measurement of E18.5 digit bone lengths found no differences between mutant and control embryos. Furthermore, when the developmental progression of cartilaginous elements was analyzed in *M280* and *M1442* embryos from E13.5-E15.5, transient development defects were not detected. These results demonstrate that *M280* and *M1442* are not required for mouse limb development. Additionally, these studies highlight the importance of experiments that manipulate enhancers *in situ* to understand their contribution to development. More broadly, our findings suggest that transcriptional regulation is complex and that no single enhancer validation method is yet capable of predicting whether an enhancer is required for development.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

An ever-expanding understanding of vertebrate limb development is desirable from at least two levels of biological inquiry: from what one may call a proximate, or developmental level, and from an ultimate or evolutionary level. At the proximal or developmental level researchers attempt to describe the molecular genetic mechanisms responsible for the specification, differentiation and growth of the limb during embryogenesis (reviewed in (Rabinowitz and Vokes, 2012) (Zeller et al., 2009) (Tickle, 2006)). The attractiveness of the limb system for molecular developmental studies derives from the complex, coordinated generation of bones, muscles, nerves and tendons along three interdependent axes: proximal-distal (shoulder/femur to digit tips); dorsal-ventral (back of hand/foot to palm); anterior-posterior [thumb (digit 1) to pinky (digit 5)]. The limb has become an ideal system within which key concepts in developmental biology are tested, confirmed, rejected and generated. For example, the concepts of diffusible signals (Rosello-Diez et al., 2011) (Cooper et al., 2011), signaling pathway crosstalk (Laufer et al., 1994) (Pajni-Underwood et al., 2007), and organizing centers embodied in the apical ectodermal ridge (AER) (Sun et al., 2002) and zone of polarizing activity (ZPA)(Hill., 2007), (Furniss et al., 2008) are interwoven into a genetic explanation for the pattering and outgrowth of the limb into three general regions: the stylopod comprising the humerus, the zeugopod comprising the radius and ulna, and the autopod comprising the bones of the wrist and the digits. Developmental mechanisms discovered in the context of limb development, and conceptual themes formulated to account for the complexity of our increasing understanding of limb development should continue to be transferrable to an understanding of the development of other structures and disease. Perhaps this is both the hope and the reason behind the consistent production of research centered on limb development, particularly in the mouse and chick.

Continued research on vertebrate limb development also appeals to questions and observations stemming from an ultimate or evolutionary level of biological inquiry (Zakany and Duboule, 2007; Shubin et al., 2009). Widely divergent vertebrate fore- and hindlimb morphologies appear in the fossil record and in extant organisms, including species with reduced or supernumerary bone numbers (Shubin et al., 2006), webbed autopods (Hockman et al., 2008) and limb bone elongations (Sanchez-Villagra and Menke, 2005; Mitgutsch et al., 2012). Exploration into how these varying morphologies appeared and disappeared in the fossil record, and how they are maintained and contribute to fitness in extant vertebrates enhances our understanding of evolutionary mechanisms. The continued study of the proximate causes of appendage development in numerous organisms has undoubtedly demonstrated that limb development in all organisms depends on the interplay of key conserved genes and signaling pathways, such as the homeobox (Hox) transcription factors, Sonic hedgehog (Shh), Bone morphogenetic proteins (Bmp) and Fibroblast growth factors (Fgf) (Shubin et al., 1997), (Rabinowitz and Vokes, 2012). Although a consensus picture of the genetic circuits underlying limb development has emerged from genetic studies in mice and chick, less is known about the genetic components responsible for the execution and regulation of these circuits.

One important form of transcriptional regulation that may simultaneously contribute to the diversity of the deployment of limb-related genetic circuits and final species-specific limb morphologies is encoded in transcriptional enhancers. Enhancers can be defined as primary sequences of genomic DNA that contain clustered transcription factor (TF) binding sites (Levine, 2010). When particular TFs or combinations of TFs bind to enhancer sequences they can act in concert with basal transcriptional machinery at a promoter to regulate the timing, location and quantity of gene expression (Banerji et al., 1981), (Visel et al., 2009b), (Ohler and Wassarman, 2010). Enhancers may be located 5', 3', or within the introns of the genes they regulate (Chandler et al., 2007; Feng et al., 2008), and the proximity of multiple TF binding

sites within enhancers confers them with modular behavior. An enhancer can be removed from its endogenous loci, linked in any orientation to a reporter gene (*lacZ*, for example), and still regulate the reporter gene transcription in a repeatable temporal and spatial pattern (Mortlock et al., 2003). Research aimed at uncovering the molecular genetic sources of divergent limb morphology, such as the role that enhancer variation may play in species-specific activation of gene circuits, serves as a bridge between the motivations underpinning questions centered on proximate and ultimate explanations of limb development.

Recent reviews have synthesized the current understanding of the importance of limb enhancers in disease etiology (and the relevance of enhancers in other forms of disease genesis and progression) (VanderMeer and Ahituv), (Sakabe et al., 2012). Other papers have hinted at the relevance of enhancer biology in explaining diverse limb morphologies characterized in paleontological and comparative embryological studies (Chen et al., 2005; Cretekos et al., 2005; Nolte et al., 2009). These papers do much to lay the conceptual and observational foundation necessary to generate excitement about the role that enhancers play in generating a diversity of limb patterns, both adaptive and maladaptive. On the other hand, functional analyses of limb enhancers (or enhancers in general) have somewhat lagged behind the enthusiasm for the topic. An appropriate way to characterize the role of a particular genetic element in development and disease is to use functional genetic methods to manipulate the element at its endogenous locus. As with genes, functional manipulation of enhancers can be achieved using conventional gene targeting techniques as well as by using somewhat more recently employed Cre-lox recombination techniques that rearrange genetic loci independent of extensive mouse embryonic stem cell handling (Herault et al., 1998), (Nagy et al., 2003), (Spitz et al., 2005) . The few limb enhancers that have been functionally manipulated have contributed greatly to our understanding of enhancer regulation generally and limb development and evolution specifically (discussed below). However, given the complexities of limb development (Taher et al., 2011a)

and limb diversity there are a great many more limb-specific enhancers to be characterized and identified. Numerous methods exist to identify potential tissue-specific enhancers; so much so that the availability of datasets composed of potential enhancers – limb and otherwise – has far outpaced careful functional genetic experimentation to characterize enhancers. In response to this situation several relevant questions come to mind: Are all methods used to identify enhancers equally valid, or are some better than others? In other words, which methods provide the least amount of false positives? Do all active enhancers in a particular tissue share a common modified-chromatin signature? If not, why not? If so, how do we filter away false positives; that is, other genetic elements that may carry the chromatin signature as well but that do not contribute to tissue-specific gene expression? Do highly conserved enhancers contribute to gene activation involved in the genesis of shared tissue elements across multiple species, while non-conserved enhancers contribute to more derived aspects of tissues? I suggest that no matter what methodology is used to identify putative enhancers, an endogenously based, functional genetic approach to enhancer characterization is necessary to address these questions. Indeed, functional genetic experiments are essential to appropriately interpret and model enhancer behaviors responsible for congenital limb malformations and the evolution of diverse limb morphologies. Therefore, below I will briefly review what is known from studies that functionally manipulated limb enhancers and discuss the methods used to identify enhancers. One method in particular – the use of Chromatin immunoprecipitation with sequencing (ChIP-seq) targeted at an enhancer-associated protein – produced a plethora of data suggesting that hundreds of enhancers contribute to region-specific development of the mouse limb. I then introduce the functional genetic approach I took to put these particular findings to the test.

Enhancer-gene proximity led to first discoveries of limb-specific enhancers

The assumption that enhancer elements were likely situated somewhere near gene transcription start sites led to the characterization of some of the earliest discovered limb-specific enhancers. The *Hoxb6* enhancer is among the earliest characterized enhancers to drive reporter gene expression in the embryonic mouse limb, although it also directs expression elsewhere including the trunk between the limbs (Schughart et al., 1991). It was determined that strong limb-specific enhancer elements lie between 1,200-3,600 base pairs (bps) upstream of *Hoxb6* because this region was consistently capable of expanding and intensifying anterior and posterior *lacZ* expression in the undifferentiated mesenchyme of early mouse limb buds when added to a *lacZ*-transgene directed by a smaller 1.2 kb upstream sequence (Schughart et al., 1991).

The limb-specific *Prx1* enhancer was identified in a manner similar to the *Hoxb6* enhancer. In transgenic mice, this *Prx1* enhancer, a 2.4 kb stretch of DNA just 5' of *Prx1,* was capable of driving *lacZ* expression that recapitulated most of the known *Prx1* limb mesoderm expression pattern from E10.5 to birth, though small differences were noted (Martin and Olson, 2000).

Neither the *Prx1* nor *Hoxb6* enhancer can direct transcription in the limb ectoderm. The AER, a portion of the distal limb ectoderm, is essential for proper limb development. The AER is required for limb outgrowth, and signals from the AER, including FGF family members, are required, via a genetic circuit, to maintain *Shh* expression in the ZPA. Two AER-active enhancers were discovered that served and serve as tools to dissect the AER's role in limb development (see below): one is associated with *retinoic acid receptor β2* (RARβ2) and the other comes from the *Msx2* locus (Mendelsohn et al., 1991; Reynolds et al., 1991; Liu et al., 1994). As was the case with the *Prx1* and *Hoxb6* enhancers, the *RARβ2* and *Msx2* AER-enhancers were relatively close to their cognate gene's transcriptional start site (located within

the 3100 bps upstream of *RARβ2;* and located within the 500 bps upstream of *Msx2*). Interestingly, *RARβ2* is not usually expressed in the AER, suggesting that when this enhancer is removed from its endogenous context it is not restrained by nearby repressor elements (Mendelsohn et al., 1991).

Endogenously manipulated limb enhancers

By uniting the tissue-specificity of limb enhancers with *Cre recombinase* and other endogenously manipulated genes (such as floxed Bmps, for example) the functional role of particular genes during limb development can be assessed. But a fuller understanding of limb development is only possible when the enhancers themselves are also manipulated in their endogenous loci; such experiments have the potential to elucidate the role of individual enhancers in the elaboration of the limb. To appropriately complement transgenic assays – which may or may not faithfully mimic the regulatory abilities of a tested enhancer at its endogenous loci –it is necessary to manipulate them within their native genomic context. The use of homologous recombination to introduce a manipulated genomic fragment into the genome of mouse ES cells, or gene targeting, has been used extensively to knockout, disrupt and modify hundreds of genes (Nagy et al., 2003). This same approach has been used, comparatively, to a limited extent to knockout or modify enhancer loci, but those loci that have been targeted tell us much about the nature of gene regulation during limb development. Moreover, two methods that employ the *Cre recombinase* technology in concert with meiotic recombination to delete, duplicate or invert portions of the mouse genome have also been used to characterize in a detailed manner the enhancer activity of one locus, the *Hoxd* cluster, in the mouse (Herault et al., 1998; Spitz et al., 2005). Insights gleaned from the functional dissection of *Hoxd* cluster regulation beg for continued research of similar intensity at other loci relevant to

limb generation. For purposes of this discussion we consider any disruption of an enhancer's endogenous locus as a functional alteration, including those caused by transgene insertions and chromosomal alterations.

Sonic Hedgehog Enhancer

While producing transgenic mice for the purpose of expressing a reporter gene in a rhombomere 4-specific manner, Sharpe *et al.* (1999) disrupted a long-range enhancer positioned approximately 850 kb from *Shh* (Lettice et al., 2002). Interestingly, later cloning efforts revealed that the insertion of the transgene concatemer resulted in the duplication of intron 5 of *Lmbr1,* the gene within which the *Shh*-regulatory sequence is located (Lettice et al., 2002). This suggested that in addition to sequence mutations/alterations, topographical changes to an enhancer could alter transcriptional output of regulated genes as well. Altered limb morphology induced by the duplicated enhancer included polydactyly, extra inverted claws and dorso-ventral patterning defects. In mutants, ectopic anterior expression of numerous genes that regulate distal limb patterning were also observed including *Shh, Ffg4/8* and *Hoxd12/13* (Sharpe et al., 1999). Such anterior misexpression of these genes had been noted before in mouse and chick models displaying limb patterning defects, including preaxial polydactyly (Riddle et al., 1993; Masuya et al., 1995; Sagai et al., 2004; Masuya et al., 2007).

The transgene-disrupted enhancer was later called the Zone of Polarizing Activity Regulatory Sequence (ZRS) because of its ability to drive *lacZ*-reporter gene expression in a manner similar to *Shh* expression in the ZPA (Lettice et al., 2003). Further confirmation that the ZRS regulated limb-specific *Shh* expression came when gene targeting was used to delete the 1.2 kb enhancer (Sagai et al., 2005). The mutant mice displayed distally truncated fore- and

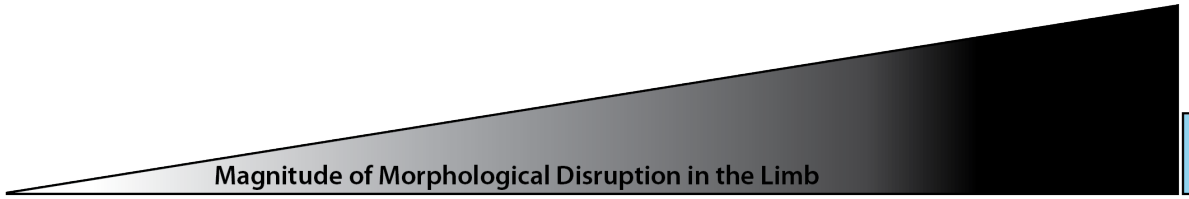
hindlimbs at E12.5 and *in situ* hybridization showed loss of *Shh* expression in developing limbs (Table 1).

The differences in phenotypes associated with the different alterations to the ZRS above hint at the complexity of enhancer-gene transcription relationships: complete deletion of the ZRS resulted in severe skeletal truncations in both fore- and hindlimbs (Sagai et al., 2005). However, a topographical change to the ZRS induced by transgene insertion resulted in an imbalanced phenotype, with hindlimbs more severely affected than forelimbs (Sharpe et al., 1999). Duplication-based mutations of the ZRS are also responsible for some reported human familial syndactylies (Wieczorek et al.; Sun et al., 2008). Moreover, numerous single nucleotide polymorphisms in this enhancer have been linked to distinct, yet related, distal limb malformations in humans (Wieczorek et al.; Furniss et al., 2008) (also, human *SHH* mutations summarized in (VanderMeer and Ahituv)). These results suggest that enhancer sequences as well as the topographical relationship between an enhancer and its regulated gene are important for proper gene transcription. The range of phenotypes associated with ZRS alterations also hints at the diversity of enhancer mutations that could potentially contribute to varied limb morphologies in extant and extinct vertebrates.

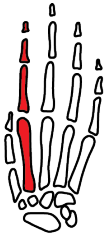
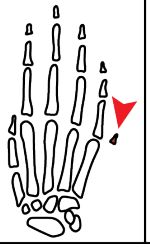
Prx1 Enhancer

The discovery and characterization of the *Prx1* enhancer was discussed previously. The ability of this enhancer to regulate limb-specific expression patterns that mimic endogenous *Prx1* suggested that its targeted deletion should phenocopy the limb malformations of *Prx1* null mice, which manifest shortened long bones in fore- and hindlimbs (Martin et al., 1995; Lu et al., 1999). However, when the limb-specific *Prx1* enhancer was deleted the limbs were indistinguishable from control littermates, suggesting that redundant *Prx1* limb enhancers

Table 1. Phenotypes associated with endogenously manipulated limb enhancers. Targeted limb enhancer deletions cause a range of phenotypes, from no alteration (*Prx1*), to moderate (*Hoxd GCR + Prox*) to severe (*Shh*). Neomorphic changes in limb patterning have been reported for a deleted (*Tcfap2α*) and replaced (*Prx1*) limb enhancer. Red indicates site of detected phenotype. See text for details and references.



Neomorphological Limb Alterations

Enhancer (gene)	*Prx1*-Enh (*Prx1*)	Proximal Gene Desert (5′ *Hoxd* genes)	Distal Gene Desert (5′ *Hoxd* genes)	GCR + Prox (5′ *Hoxd* genes)	HLEA (*Tbx4*)	*Gremlin*-GCR (*Gremlin*)	ZRS (*Shh*)	DCE (*Tcfap2a*)	*Prx1*-Enh^{Bat} (*Prx1*)
Size: deleted sequence	1 kb	300 kb	290 kb	240 kb	1.1 kb	25 kb	1.2 kb	530 bp	1 kb; replacment
Enhancer-gene distance	Proximal to promoter	~250 kb	~550 kb	~50 kb	~10 kb	~45 kb	~850 kb	5th intron	Proximal to promoter
Phenotype									

compensate for its loss (Cretekos et al., 2008). Interestingly, when gene targeting was used to replace the mouse *Prx1*-limb enhancer with the orthologous bat enhancer (from *Carollia perspicillata*) the homozygous bat-enhancer-mice develop slightly longer forearms and express 1.7 fold higher amounts of *Prx1* than control littermates (Cretekos et al., 2008). These results demonstrate that although redundant enhancers regulate *Prx1* limb expression, minor sequence variation in a single enhancer can account for some of the length differences observed in mouse and bat forelimb morphology [~93.5% sequence similarity between mouse and bat in the two core conserved regions of this enhancer (Martin and Olson, 2000; Cretekos et al., 2008)].

Tcfap2α Enhancer

Like *Prx1*, *Tcfap2a* is required for mouse craniofacial and limb development. Mice lacking both copies of *Tcfap2a* lose their cranial neural crest cell derivatives, display midline clefting and possess malformed radii (Schorle et al., 1996; Zhang et al., 1996; Nottoli et al., 1998; Brewer et al., 2004). Zhang *et al.* (2003) identified an enhancer element in the fifth intron of *Tcfap2a* that could direct *lacZ* expression to the frontal nasal process and limb bud mesenchyme when linked to the *Tcfap2a* minimal promoter region (Zhang and Williams, 2003). Further refinement of this enhancer identified the downstream conserved element (DCE), which is highly conserved in all vertebrates, indicating that the DCE performs an essential role in regulating *Tcfap2* expression during vertebrate appendage development (Donner and Williams, 2006; Feng et al., 2008).

To functionally test this hypothesis, Feng *et al.* (2008) used gene targeting to remove 530 bps from the *Tcfap2a*-fifth intron, which included the entire DCE. Concordant with the known role of the DCE in regulating limb-specific *Tcfap2a* transcription, mice bearing a homozygous deletion of the DCE invariably manifested a patterning defect in the forelimbs, but

not in the hindlimbs. Forelimbs of DCE null mice possessed an extra posterior digit, concordant with reduced *Tcfap2a* expression in mutant limb buds (Feng et al., 2008). Thus, contrary to the results obtained from the *Prx1*-enhancer knockout mice, the deletion of *Tcfap2a's* fifth intron-enhancer demonstrates that single enhancers can play important roles during limb development. On the other hand, the *Tcfap2a*-enhancer null mice did not recapitulate the limb phenotypes observed in mice harboring *Tcfap2a -/-* alleles (Schorle et al., 1996; Zhang et al., 1996; Nottoli et al., 1998; Brewer et al., 2004) thus the action of partially redundant *Tcfap2a*-limb enhancers at this locus seems likely.

Tbx4 Enhancer

Several genes are differentially expressed in the fore- and hindlimb, including *Tbx4* and *Tbx5*. *Tbx4* is preferentially expressed in the hindlimb (Chapman et al., 1996); although mice lacking this gene die in mid-gestation, hindlimb development was still observed to have been adversely affected (Naiche and Papaioannou, 2003).

To identify hindlimb-specific enhancer elements Menke *et al.* (2008) inserted an IRES-βgeo reporter construct into the 3'-UTR of *Tbx4* in three separate BAC transgenes that spanned varying lengths of the *Tbx4* locus (Menke et al., 2008). Eventually, two core enhancer elements residing in the BACs were characterized via their ability to drive *lacZ* expression: hindlimb enhancer A (HLEA) is 1073 bps, and hindlimb enhancer B (HLEB) is 654 bps. Interestingly, both HLEA and HLEB were capable of driving *lacZ* expression in a medial forelimb domain in transgenic mice at E10.5-11.5. Thus their ability to direct forelimb gene expression is likely suppressed at their endogenous locus by additional insulating regulatory elements.

Because removal of HLEA from the IRES-βgeo-containing BAC had the greatest impact on depleted βgeo staining, gene targeting was used to remove this enhancer from the mouse

genome to test its function during development. Whole mount *in situ* hybridization in E10.5 embryos showed that *Tbx4* expression was reduced in mutant hindlimbs. These expression differences presaged the morphological changes observed in adult mice, which included shortening of pelvic, femur, patella and tibia bones along at least one of their axes. Additionally, hindlimb digits 1, 3 and 5 were significantly smaller when compared to those of control littermates (Menke et al., 2008). Again, these results demonstrate that a single enhancer can harbor regulatory information necessary for proper limb development and that disruption of a single enhancer could potentially provide varied morphological substrates for evolution. However, the hindlimb phenotype in HLEA null mice is less severe than that observed in mice with a temporally guided, limb-specific deletion of the *Tbx4* gene indicating again that other enhancers partially compensate for the loss of HLEA (likely HLEB in this case) (Naiche and Papaioannou, 2007).

Hoxd Cluster

The *Hoxd* cluster genes, which along with *Hoxa* genes are required for autopod development (Davis et al., 1995; Fromental-Ramain et al., 1996; Zakany and Duboule, 1996; Zakany et al., 1997), are regulated by a wide array of enhancers, located both on the 5'- or centromeric and 3'- or telomeric sides of the cluster. Functional tests employing numerous chromosomal inversions and deletions using two *Cre/loxP*-based methods, TAMERE and STRING (Herault et al., 1998; Spitz et al., 2005), aided researchers in illuminating the multifaceted regulation of the *Hoxd* cluster.

Several mouse lines carrying human or mouse *Hoxd*-cluster-containing transgenes, along with the mapping of an X-ray induced inversion that generated mice lacking nearly the entire

ulna, indicated that the centromeric side of the cluster housed its limb-specific enhancers (Davisson and Cattanach, 1990; Herault et al., 1997; Peichel et al., 1997; Spitz et al., 2001). The initial centromeric enhancer region, some 240 kbps upstream from the cluster, was later shown to contribute to distal limb reporter gene expression in BAC transgenic mice (Spitz et al., 2003; Gonzalez et al., 2007). The region spanned roughly 40 kb of strongly conserved sequence and was termed the Global Control Region (GCR) because of its ability to regulate limb expression of numerous genes in the vicinity (e.g. *Lnp*). It was later shown that another centromeric enhancer region, called Prox, also contributed to the proper spatial expression of 5' *Hoxd* genes, such as *Hoxd13* (Gonzalez *et al.*, 2007). Specific elements within the GCR and Prox, called the CsB and CsC, respectively, directed *lacZ* expression in the autopods of transgenic mice. Interestingly, in a large transgene construct containing both the entire GCR and Prox regions, but not the intervening sequence, the distal- and anterior-most expression pattern of *Hoxd* genes was not recapitulated, suggesting that the relative positioning of these two elements – their genomic architecture – is also important for faithful endogenous gene expression (Gonzalez et al., 2007).

Further characterization of the GCR's regulatory role crystalized when Spitz *et al.* (2005) used the STRING method to invert a chromosomal region spanning the *Hoxd* cluster that resulted in a 3 Mb separation between *Hoxd1-10* and *Hoxd11-13. In situ* hybridization against *Hoxd10* showed that its expression was lost in the distal limb buds but retained in the proximal limb buds; the opposite was observed for *Hoxd11.* Therefore, the GCR, which maintained its association with *Hoxd11* throughout the inversion, continued to regulate its distal expression. And *Hoxd10*, though out of range of the GCR, maintained its proximal expression domain, presumably due to another enhancer region on the telomeric end of the cluster called the Early Limb Control region (ELCR) (Zakany et al., 2004; Spitz et al., 2005).

The TAMERE method was used to produce a series of deletions and expansions of the internal *Hoxd* cluster and permitted a better characterization of the ELCR's role in *Hoxd*-cluster limb expression (Herault et al., 1998). For example, internal *Hoxd* cluster deletions placed 5' *Hoxd* genes closer to the ELCR. These genes – such as *Hoxd13*, which is usually expressed in autopodal posterior margins after E10 – were now expressed in the anterior and posterior margins of E9-9.5 limb buds. *Hoxd13's* topographical shift towards the telomeric side of the cluster put it under the control of the ELCR, which is capable of directing earlier, more anterior expression of *Hoxd* genes (Tarchini and Duboule, 2006).

The GCR and ELCR are not the only stretches of enhancers required for proper temporal and spatial expression of the *Hoxd* genes in the limb. When a gene desert centromeric to the GCR was inverted using STRING (Spitz et al., 2005), the transcriptional output of *Hoxd10-13* was reduced relative to control limbs, suggesting that important topographical interactions between additional enhancers and *Hoxd* genes had been disrupted. Four discrete enhancers within the centromeric gene desert – whose regulatory potential was predicted based on chromosome conformation capture assays (both 3C and 4C), sequence conservation, and enhancer-related chromatin marks (discussed below) – directed *lacZ* expression in the limbs of transgenic mice (Montavon et al., 2011).

The various chromosomal rearrangements discussed above illuminated the location of limb-specific regulatory regions relative to the *Hoxd* cluster and provided further evidence that the distance between enhancers and genes can influence gene expression and ultimate limb patterning.

To assess the importance of specific enhancers or enhancer-containing regions within the *Hoxd* cluster topography, Cre-mediated recombination was used to delete various enhancer elements. Interestingly, the numerous enhancer elements spread throughout the approximately 830 kb large gene desert centromeric to the *Hoxd* cluster seem to be partially complimentary,

with no single enhancer-region deletion resulting in drastic multi-digit malformations. Still, minor phenotypes are associated with single enhancer-region deletions; for example deletion of a 240 kb region spanning GCR and Prox results in the absence of a single phalange in digit 5. On the other hand, deletions that span several enhancer regions lead to increasingly severe digit phenotypes. For example, a deletion spanning the entire 600 kb gene desert results in phalange loss in all digits, while a 830 kb deletion that removes the entire characterized centromeric enhancer expanse (gene desert + GCR + Prox) results in a severely stunted digit array with reduced bone growth in newborn mice; this phenotype resembles that of homozygous *Hoxd8-13* null mice (Montavon et al., 2011).

The surprisingly subtle nature of the phenotypes associated with deletions of vast enhancer regions indicates that the autopodal transcription of *Hoxd* genes is well buffered against perturbations that may affect the activity of any single enhancer element. These observations also suggest that isolated non-topographical sequence variations within any given enhancer element in this region are unlikely to be responsible for large differences in autopod morphology between species. Rather an accumulation of numerous sequence changes across the centromeric enhancer interval would likely be required to contribute to striking differences in vertebrate autopod morphology. It has been demonstrated in transgenic mice, however, that a Greater Horseshoe Bat GCR-ortholog is capable of driving *lacZ* expression in expanded zeugopod and stylopod domains in addition to the expected autopod domain (Ray and Capecchi, 2008) (see also (Chen et al., 2005)). This indicates that species-specific regulatory sequences reside in the bat GCR, though they were not mapped to the GCR-CsB enhancer or any other GCR region. In contrast to, or in concert with enhancer sequence variation, the above discussion indicates that chromosomal rearrangements, even relatively minor insertions and deletions within and around the *Hoxd* locus (Gonzalez et al., 2007), could greatly impact the temporal and spatial expression of *Hoxd* genes and thus the ultimate morphology of the limb.

Gremlin Global Control Region

Work on the limb-specific *Gremlin*-GCR suggests that global control regions may not just be a unique regulatory feature of the collinearly expressed *Hox* gene loci. The GCR that regulates autopodal expression of *Gremlin* and the nearby gene, *Formin*, was discovered in a circuitous fashion. The phenotypic effects of disrupting this GCR were initially attributed to the disruption of *Formin*, approximately 40 kb downstream of *Gremlin*, because of its relationship to mapped lesions that made up the *limb deformity* (*ld*) mutation (Maas et al., 1994; Wang et al., 1997; Zeller et al., 1999). However, two other *ld* alleles identified from phenotypic screens were then discovered to disrupt *Gremlin*, placing *Gremlin* in the same *ld* complementation group as *Formin.* Moreover, *Formin* null mice do not manifest limb defects but do present partial renal agenesis, whereas *Gremlin* null mice possess malformed limbs due to AER disruption. This led Zuniga *et al.* (2004) to hypothesize that the *ld* lesions in *Formin* were actually disrupting enhancer elements responsible for limb-specific expression of *Gremlin (Zuniga et al., 2004)*.

Confirmation of this hypothesis came when gene targeting was used to delete exon 10 of Formin, completely disrupting its C-terminal domain. While mice homozygous for the *Formin*-exon 10 deletion have normal limbs, another line of mice possessing a deletion from *Formin* exons 10-24 phenocopied the *ld* limb defects. *Gremlin* expression was absent in the autopod of these mice. Therefore, it was concluded that regulatory elements internal to the *Formin* exon 10-24 interval regulated limb expression of *Gremlin (Zuniga et al., 2004).* Further confirmation of this conclusion came when a BAC transgene that harbored *lacZ* in the *Gremlin* locus recapitulated endogenous autopod *Gremlin* expression only when the *Formin* exon 19-24 interval was present; when the exon 19-24 interval was removed from the BAC *lacZ* expression was abolished in the developing limb (Zuniga et al., 2004).

Methods used to identify limb enhancers

In the post-genomic era, cross-species sequence conservation is an integral part of enhancer identification, either as a starting place, as when an enhancer is predicted based solely on sequence comparisons (Pennacchio and Rubin, 2001; Haeussler and Joly, 2011), or as a validation of other methods, as when a region bound by a TF in a ChIP-seq experiment is also highly conserved between species (Visel et al., 2009a). An increasingly diverse array of sequence comparison assumptions and algorithms are augmenting the detection of sequences encoding functional enhancer activity. For example, the assumption that genes and their associated regulatory milieu need to be inherited and selected together prompted Ahituv *et al.* (2005) to utilize cross-species syntenic regions to identify conserved non-coding sequences that could potentially regulate developmental genes (Ahituv et al., 2005). Another study demonstrated that orthologous yet non-conserved (defined as below 70% identity over 100 bp), non-coding sequences, called covert elements, could be detected when the ordering and distance between TF binding sites within these elements was taken into account (Taher et al., 2011b). Taher *et al.* (2011) showed that 7/8 covert elements from human and zebrafish discovered using their method drove similar reporter gene expression in a zebrafish transgenic assay. An AER-specific enhancer near the *Dach1* locus, which is itself expressed in the AER (Caubit et al., 1999), was identified using yet another sequence comparison algorithm called Gumby. The *Dach1*-AER enhancer was one among numerous putative conserved enhancers identified from sequence comparisons of recently diverged species (e.g. human and mouse vs. human and frog) characterized by extremely low Gumby p-values (Prabhakar et al., 2006). Interestingly, another method, focusing on sequences conserved across vast phylogenetic distances, from human to fish (*fugu* in this case), had also resulted in identification of a limb-specific enhancer near the

Dach1 gene (Nobrega et al., 2003). Despite the somewhat complex complimentary and divergent natures of these sequence-based enhancer discovery methods one outcome of their application is clear: there are now large quantities of potential enhancers identified near genes expressed throughout limb development. Their role in development and evolution awaits analysis via transgenic and functional genetic methods.

Further bioinformatic and evolutionary approaches can be applied to the criterion of sequence conservation to identify enhancers that may drive transgenic *lacZ* expression in unique, species-specific ways. For example, Prabhakar *et al.* (2008) first identified non-coding regions of the genome that were highly conserved between all sequenced terrestrial vertebrates. This data set was further filtered using a statistical test that identified human sequences that had rapidly diverged from other vertebrates. The authors then focused on a 546 bp sequence, human-accelerated-conserved-noncoding-sequence 1 (HACNS1), which was the fastest evolving human-specific sequence in the last 6 million years, the time since the last human-chimpanzee common ancestor. When HACNS1 was ligated to a minimal promoter and *lacZ* it drove expression in the E11.5 mouse limb bud, including the distal tip, suggesting it may have played a role in shaping unique human limb morphologies (Prabhakar et al., 2008). It would be interesting to functionally test the role of HACNS1 during limb development by either deleting it or replacing the homologous sequence in mice with those from the human and chimpanzee.

In addition to conventional sequence-based approaches used to identify enhancers are those methodologies that take advantage of more recent chromatin immunoprecipitation (ChIP) and sequencing technologies. As an example, Vokes *et al.* (2008) targeted the ubiquitously expressed *Rosa 26* locus with a floxed-stop cassette and a flag-tagged *Gli3*-repressor (*Gli3R*) cDNA. Therefore, upon limb-specific, Cre-mediated recombination the stop cassette was removed and the tagged *Gli3R* was expressed. Genomic regions interacting with Gli3R, termed Gli binding regions (GBRs), were then identified via ChIP-on-chip. The regulatory potential of

several GBR enhancers was then tested by scoring their ability to drive *lacZ* expression in transgenic mice and comparing the *lacZ* expression pattern with that from genes associated with the GBRs. In this manner Gli-associated enhancers were shown to drive limb-specific *lacZ* expression patterns mimicking portions of the limb-specific *Gli1, Blimp1* and *Gremlin* expression domains (Vokes et al., 2008).

In a similar vein, Kouwenhoven *et al.* (2010) coupled their knowledge that heterozygous mutations in *p63* correlate with split hand/foot malformation (SHFM) with ChIP-seq methodologies to identify binding sites of p63 throughout the mouse genome. Then using a set of bioinformatic criteria, such as proximity of binding sites to the *Dlx 5/6*-containing SHFM 7q21 interval, the researchers identified putative p63-regulated enhancers. One enhancer, p63-binding site (BS) 1 was shown to drive reporter gene expression in both zebrafish and mouse AERs. Importantly, when the p63 binding site was mutated in the BS1-transgene construct, no reporter gene expression was noted in mouse AERs (Kouwenhoven et al., 2010). This work demonstrates that methodologies based on single transcription factors can be used to identify enhancers with limb-specific activity.

Recently, ChIP-seq was used to provide a catalog of tissue-specific enhancers available to researchers for further characterization or manipulation. This resource is found on the Vista Enhancer Browser (http://enhancer.lbl.gov) (Visel et al., 2007b). Tissue-specific enhancers made available on this site are first tested for their ability to direct *lacZ* expression in transgenic mice; images of the mice along with the sequence and chromosomal location of the enhancer are provided. Numerous enhancers have been tested that were identified based on sequence conservation alone (Nobrega et al., 2003; Pennacchio et al., 2006; Visel et al., 2007a; Visel et al., 2008) [also reviewed in (Pennacchio and Rubin, 2001)]. In 2009, enhancers from a ChIP-seq-based data set were added to the Vista Browser (Visel et al., 2009a). Limb-specific enhancers in this data set were discovered by homogenizing E11.5 mouse limbs and sequencing

the DNA crosslinked to antibody bound-p300. p300 is a general transcriptional co-activator with histone acetyltransferase activity associated with enhancers and active transcription (Heintzman et al., 2007; Xi et al., 2007) (also reviewed in (Chan and La Thangue, 2001)). A subset of the p300-associated enhancers derived from limb tissue were tested for their ability to reproducibly regulate transgenic *lacZ* expression in E11.5 mice. However, to date none of the transgenically examined p300 limb-specific enhancers have been tested in an *in vivo* functional genetic context.

Related to enhancer-p300 associations are a number of other chromatin marks used to identify putative enhancers via ChIP-seq methods, including H3K4me1 and H3K27ac (Heintzman et al., 2007; Hon et al., 2009; Ong and Corces, 2011; Cotney et al., 2012). The rapidly accumulating data on whole genome chromatin signatures for enhancers can be creatively combined with uniquely tailored bioinformatic screens to identify novel types of regulatory elements. For example, Birnbaum *et al.* (2012) took advantage of 25 available ChIP-seq datasets that recorded the chromatin-related marks of active enhancers. The authors then took these datasets through a series of screens to identify sequences that overlapped with the previously discussed p300 dataset and with exons (rather than introns or intergenic sequences) of genes that were not themselves expressed in limbs but were near genes that were, such as *Dlx5* and *6*. In this manner Birnbaum *et al.* (2012) demonstrated that four of these sequence types, termed enhancer-exons (eExons), regulated a *lacZ* expression in the E11.5 mouse limb (Birnbaum RY, 2012).

The usefulness of this approach was further confirmed by employing chromosome conformation capture (3C), another technique that is increasingly used to decipher enhancer-enhancer and enhancer-gene interactions. In this case, a 3C assay confirmed that one of the eExons, *Dyn-eExon15*, associated with the promoters of two genes some 900 kb away, *Dlx 5* and *6*, both of which are related to SHFM (Birnbaum RY, 2012). Both 3C and 4C (another

flavor of 3C, used when only one potentially interactive sequence is known) were also used to better understand the complex interactions between the *Hoxd* cluster with both the extensive centromeric (5') and telomeric (3') enhancer regions; these experiments showed that multiple enhancer elements, including the GCR, Prox and enhancers within the centromeric gene desert come together to regulate 5'-*Hoxd* gene members, such as *Hoxd13* (Montavon et al., 2011).

Another attempt to catalog and map enhancers throughout the entire mouse genome resulted in the creation of a regulatory or enhancer sensor. The regulatory sensor (called SBlac) is composed of, 5'-to-3', a single *loxP* site, the human *β-globin* promoter and the *lacZ* reporter, all flanked by the Sleeping Beauty transposon inverted repeats (Ruf et al.). The *β-globin* promoter has no activity on its own, and therefore can only aid in *lacZ* transcription when next to enhancer elements. Offspring generated from mice harboring both a spermatid-specific Sleeping Beauty transposase transgene and the SBlac transgene sensor therefore have SBlac insertions in unique locations throughout the genome. The SBlac is capable of "reading" the regulatory potential at its unique insertion site. Ninety-eight out of 165 (60%) transgenic mice scored for *lacZ* expression showed tissue-specific expression patterns, including some in the E11.5 limb. Interestingly, SBlac transgenes and Vista-enhancers (deposited on the Vista Enhancer Browser (Visel et al., 2007b)) located near each other regulated similar expression patterns, with SBlac-regulated expression patterns always being more restricted. These findings accord well with those from the functional studies already discussed at the *Prx1, Hoxd* and *Tcfap2a* loci that enhancer activity at endogenous loci is often different from that characterized in ectopic transgenic experiments.

The single *loxP* site in each SBlac sensor can also be use to modify the genome around enhancers using TAMERE or STRING (Herault et al., 1998; Spitz et al., 2005); in this manner the location of specific enhancers can be ascertained. For example, Ruf *et al.* (2011) used local transposition to place two separate SBlac sensors adjacent to *Myc* in two different mouse lines.

Each sensor recapitulated portions of *Myc*'s endogenous expression pattern. However, when the lines were crossed and TAMERE was used to delete the 830 kb interval between the SBlac sensors the remaining sensor no longer directed a *Myc*-like *lacZ* expression, indicating that the enhancers regulating that portion of *Myc*'s expression had been deleted (Ruf et al.). Similar experiments could be used to both identify and, importantly, functionally manipulate limb-specific enhancers at their endogenous loci, including enhancers that were identified in two separate assays, such as the Vista-enhancers mentioned previously.

Many of the techniques described above generate vast lists of potential tissue-specific enhancers. And although a selection of enhancers pulled out of each type of screen were subjected to additional transgenic assays in either mice or zebrafish to better characterize their regulatory roles, none of them, except for a few enhancer loci in the SBlac system, were manipulated in a functional genetic manner. It is worth noting that many of the putative enhancers collected in these assays are cross-referenced to the Vista-enhancer database; presumably an enhancer found it two or more datasets is less likely to be a false positive. Recall that this database includes enhancers that were identified because of their extreme conservation as well as enhancers that proved to associate with p300 in a ChIP-seq experiment. However, to date, none of the Vista-p300-associated enhancers have themselves been deleted from the mouse genome. Triangulation of independent results as a measure of confidence in methodologies is logically sound, but at some point I believe it is important to test the validity of enhancer-discovery methods on their own merits by using gene targeting to delete a subset of the putative enhancers they produce.

Conceptual themes in enhancer biology derived from functional genetic experiments of limb-specific enhancers

After surveying the available research on functionally manipulated, limb-specific enhancers it is useful to draw attention to four conceptual themes that emerge from the disparate motivations, methods and findings connected to each enhancer study. No priority is intended in the ordering of their presentation. These conceptual themes can serve as general assumptions and observations to be refined, refuted or expanded by further research on limb enhancers.

Firstly, it is commonly appreciated that transgenic assays – wherein individual enhancers are used to drive reporter gene expression in non-endogenous integration sites – do not indicate the necessity or even the *in vivo* function of the tested enhancer. When this conceptual appreciation is acted on via the use of functional genetic approaches to modify (often delete) the enhancer at its endogenous locus the results can be stunningly telling. For example, enhancer deletions informed us of the apparent complete redundancy of the *Prx1* enhancer (Cretekos et al., 2008) versus the morphological dependency of the limbs on the *Tbx4* and *Tcfap2a* enhancers (Feng et al., 2008; Menke et al., 2008). No matter the final outcome of enhancer deletion experiments one commonality among all of the studies is the observation that phenotypes do not perfectly correlate with the entire spatial expression domain of enhancers as assayed via transgenic reporter mice. This indicates that in all cases enhancers behave differently – likely due to local interactions with other enhancers, repressors and insulators – when in endogenous or ectopic contexts (for examples see (Cretekos et al., 2008; Feng et al., 2008; Menke et al., 2008; Montavon et al., 2011). Thus, despite the operational definition of enhancers as autonomous, modular regulatory units (Visel et al., 2009b; Levine, 2010; Ohler and Wassarman, 2010), in reality they exert their influences in a context-dependent manner. This point may seem trivial, but it is only after carrying out functional, endogenous experiments on enhancer loci in order to parse their ectopic/endogenous behaviors that additional hypothesis about transcriptional regulation can be tested: What elements partially (or, more rarely, fully) compensate for this enhancer? What regulatory sequences work in concert with or against this

enhancer? How common are transcriptional enhancer networks? (Again see discussions in (Cretekos et al., 2008; Feng et al., 2008; Menke et al., 2008; Montavon et al., 2011)). Moreover, it has been elegantly argued that meaningful discussions about the homology or novelty of limb structures must take into account the regulatory mechanisms responsible for limb development (Woltering and Duboule, 2010); such mechanisms can only be determined by carrying out functional genetic experiments.

The second general observation gleaned from the functional genetic experiments above is that although the regulation of transcription by enhancers seems to be structured in a network or web-like fashion, complete redundancy, where the absence of one enhancer is completely compensated for by another (nearly always un-characterized) enhancer, is rare. Indeed, complete compensation by an as yet uncharacterized enhancer was seen only in the absence of the *Prx1* enhancer (Cretekos et al., 2008). This observation held true even for deleted chromosomal regions near the *Hoxd* cluster that presumably contain numerous enhancer elements – deletions of hundreds of base pairs did not completely compromise autopod morphology, implying the action of a regulatory network with compensatory effects. Yet the absence of numerous enhancer elements did result in observable, even if subtle, anatomical alterations, illuminating the unique role of individual *Hoxd* regulatory regions (Montavon et al., 2011). In should be mentioned too that in cases where no phenotype is scored in the absence of an enhancer, such as the *Prx1*-enhancer, this does not strictly mean that there is no phenotype – it only suggests that no phenotype was detected. Again, if indeed gradualism is the norm in evolutionary processes, and if that gradualism is at least partially encoded in enhancer elements then we may expect the phenotypes in enhancer-deleted mouse lines to be subtle, or, perhaps, only detectable under certain conditions that mimic natural environments.

The third generalization stems from the observation of compensatory enhancer networks discussed above and from the phenotypic subtlety of enhancer knockout mice relative to mice

harboring homozygous deletions for the genes they regulate. For example, when *Tbx4* is deleted at early time points in the hindlimb, development of the limb bud fails to proceed (Naiche and Papaioannou, 2007). In contrast, HLEA (*Tbx4* enhancer) null mice possess a milder phenotype, exhibiting slightly shortened (though statistically significant) digits, for example (Menke et al., 2008). Similarly, the multiple individual enhancer-region deletions (such as the GCR) surrounding the *Hoxd* locus are manifested in mice as modestly malformed autopods with missing or altered phalanges. Only after the deleting of multiple *Hoxd* enhancer-regions simultaneously do the autopods of mice reach the level of malformation seen in *Hoxd8-13* null mice (Montavon et al., 2011). There are exceptions of course; the *Shh-ZRS* null mice lack distal limb elements altogether (Sagai et al., 2005), but this is just an extreme example of the general observation that individual enhancers, despite redundancy, do contribute to the elaboration of a functioning limb, often in a piecemeal and subtle fashion (Guenther et al., 2008). Such observations suggest that mutations to individual enhancers, most strictly those leading to deletions and expansions, would at least lead to modest morphological variation, if not large. And these subtle variations could obviously serve as substrates for evolutionary phenomena, such as natural selection or genetic drift, for example. Developmental biologists pursuing functional genetic experiments to evaluate the role of enhancers need to be prepared to analyze subtle phenotypes (relative to single and bi-genic knockout experiments) and to advocate the relevance of these phenotypes for understanding the molecular genetic basis of evolution (Baguna and Garcia-Fernandez, 2003), where genetic causality may more often than not be allied with subtle variation rather than the striking deformities accompanying gene-based disease research.

Lastly, large chromosomal inversions and deletions at the *Hoxd* locus highlighted the importance of topography between enhancer-enhancer and enhancer-gene relationships in promoting proper transcription. If the distance between an enhancer and its associated gene is

too great or to small, proper regulation is lost, as when *Hoxd10* was no longer expressed in distal limb buds after an inversion had placed it out of range of the distal-acting enhancers (Spitz et al., 2005). In another experimental approach, 4C-chromosome capture was used to directly detect long-range interactions between multiple enhancers involved in *Hoxd* gene transcription (Montavon et al., 2011). Enhancer duplications, such as those recorded for the *Shh-ZRS* and *Bmp2* enhancer (Sharpe et al., 1999; Lettice et al., 2002; Dathe et al., 2009), also suggest that rearrangements in topography may be as important as the enhancer sequences themselves in dictating appropriate transcription. Directly testing the importance of topography at loci other than the *Hoxd* cluster will be necessary to fine tune our understanding of the contribution of topographical variation to gene transcription in normal, disease and species-specific contexts.

Functional analysis of p300-associated, limb-specific enhancers in the mouse

Of the numerous limb-specific enhancers discussed above it is interesting to note that a large proportion of them were in one sense or another identified based on two criteria: proximity to a gene of interest and sequence conservation (there are exceptions of course, the *ZRS* that regulates *Shh,* for example). An exciting possibility that results from screens that are gene-blind, such as those that focus on chromatin marks, is that enhancers relevant to limb development and evolution may be discovered regardless of their association to a particular gene. The Vista-ChIP-seq dataset produced by Visel *et al.* (2009), for example, identifies enhancers based on their association with p300, not on their proximity to any gene. Because p300 possesses histone acetyltransferase activity its presence at a particular locus is considered a proxy for the H3K27ac chromatin modification, which itself is considered a signature of active enhancer regulated transcription. In this sense, the p300 ChIP-seq dataset has a biological justification

beyond sequence conservation or gene proximity to be trusted as a method of active enhancer discovery. Therefore, to understand the roles of endogenous enhancers during vertebrate development I deleted two Vista-p300-associated enhancers from different genomic loci known to be active in the mouse embryo limb bud (enhancer.lbl.gov). Importantly, the enhancers were selected because of unique *lacZ* expression patterns driven by human enhancer sequences in E11.5 mice – images of which are available on the VEB. Thus, the outcome of my enhancer targeting experiments would permit me to simultaneously assess the utility of the p300 ChIP-seq dataset in identifying required limb-specific enhancers and to assess my findings in light of the results and concepts derived from the above discussed functional experiments. This is the first study to explore the consequences of deleting enhancers selected solely because of their association with p300 in a particular tissue (the limb). Thus, broadly it will address the value of the p300 ChIP-seq dataset in identifying enhancers required for development. I hypothesized that limb development and ultimate limb morphology would be disrupted in mice harboring deletions of p300-associated, limb-specific enhancers.

Transgenic mice: Selection of putative transcriptional enhancers for functional genetic manipulation

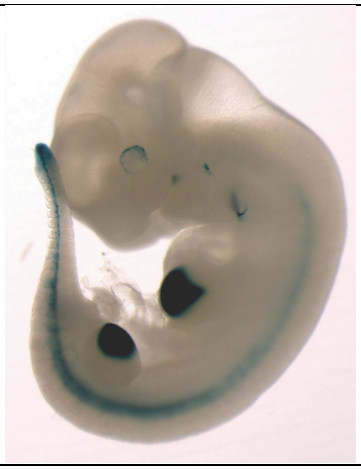
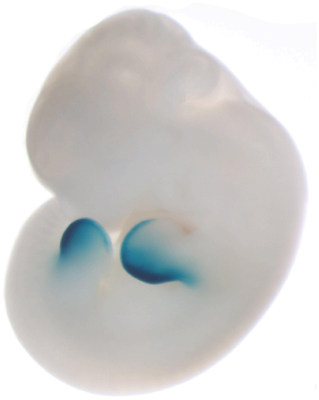
The transcriptional enhancers evaluated in this project were derived from bioinformatic analyses of vertebrate genomes and functional genetic studies conducted in the mouse (discussed below). These earlier studies relied on the human genome as the base or reference sequence for the bioinformatic and/or functional genetic approaches. I will describe how the enhancers that I manipulated were cloned and constructed into transgenic vectors and gene targeting constructs based on the findings of this previous research, starting with *M280*, and then discussing *M1442*.

M280 enhancer: Identification and validation

In 2004, Bejerano *et al.* identified 481 non-coding segments of the human genome, termed ultraconserved elements, which are perfectly conserved (100% sequence identity) between the human, mouse and rat genomes across at least 200 bps. The extreme conservation of these non-coding sequences suggested they might play an essential role in regulating gene transcription during vertebrate development. To test this, Pennacchio *et al.* (2006) made a collection of transgenic mice harboring different transgenes composed of an ultraconserved element with ± ≈ 500 bp of flanking sequence linked to the *lacZ* reporter gene (Kothary et al., 1989; Pennacchio et al., 2006). The ultraconserved segments and their flanking sequences were obtained from human DNA. Transgenic mouse embryos were collected and scored for positive *lacZ* expression at E11.5. Among the numerous *lacZ* positive embryos obtained from this experiment were four that carried the same transgene and expressed *lacZ* in a consistent, distinctive pattern in the middle of E11.5 limb buds (Table 2). Because each embryo harbored an independent transgene insertion event it was concluded that the ultraconserved element, called *hs280*, was responsible for the reproducible *lacZ* expression pattern in the limb buds and not some other genomic regulatory elements nearby to the independent insertion events of the *hs280-lacZ* transgenes (*H280*-tg).

In all four *H280*-tg E11.5 mouse embryos analyzed (Pennacchio et al., 2006), the *lacZ* expression pattern forms a somewhat equilateral triangular or wedge-shaped pattern in the middle of the fore- and hindlimb buds. The base of this triangular expression pattern runs along the distal tip of the bud, extending about half way up the distal-anterior edge and half way down the distal-posterior edge of the limb bud. The equilateral sides of the expression pattern extend from the anterior-most and posterior-most points of the distal margin of expression,

Table 2. Highly conserved and p300-associated enhancer sequences obtained from the VISTA Enhancer Browser (VEB; enhancer.lbl.gov). Blue text refers to human genomic sequence data/information taken directly from the VEB and Bejerano, G. *et al.* (2004) and Visel, A. *et al.* (2009). Black text refers to mouse genomic sequence; mouse enhancer sequences deleted in this study are in bold. Images of *hs280* and *hs1442* transgenic mice taken directly from the VEB (enhancer.lbl.gov; Visel *et al.* 2007b).

	VISTA-*hs280*	**VISTA-*hs1442***
Human enhancer drives *lacZ* expression in E11.5 transgenic mouse limb buds		
Number of *lacZ*-positive embryos/ transgene-positive embryos	**4/4**	**10/11**
Method of discovery	**Ultraconservation + p300-association**	**p300-association**
Overlapping position in mouse genome (mm9)	Chr. 4: 117,063,735-117,065,404	Chr. 1: 165,827,290-165,830,012
Name of mouse ortholog	***M280***	***M1442***
Total length	1,676 bp	2,723 bp
Both enhancers are intronic: Host gene	*Rnf220* (2nd intron)	*Kifap3* (19th intron)
Percent identity of mouse-on-human BLAT search (UCSC Genome Browser; mm9-on-hg19)	94.9% (across entire length of enhancer)	91.9% (across 1,440 bp of total length)

respectively, and meet at a point – analogously the tip of the triangle, pointing dorsally – in the center of the limb bud. This well-defined expression pattern corresponds to the future location of digits II-III and potentially to elements of the central wrist and ankle. Three of four embryos also displayed *lacZ* expression in the first pharyngeal arch and the same proportion scored positive for *lacZ* in the neural tube. Two of four embryos scored positive for *lacZ* expression in the tail bud. Images of *H280*-tg mice and others from Pennacchio *et al.* (2006) are freely accessible on the Vista Enhancer Browser (VEB; enhancer.lbl.gov) (Pennacchio et al., 2006; Visel et al., 2007b). The observation that the *H280*-tg expression pattern overlapped with limb mesenchyme that would eventually give rise to digits II-III – a region of the developing limb that is rarely studied independent of its relationship to the molecular processes set in motion by the well-known expression of *Sonic Hedgehog* (*Shh*) in the posterior limb bud – and the fact that four-of-four transgenic embryos displayed a consistent limb expression pattern led to my decision to further investigate the role of *hs280,* or its orthologs, in limb development.

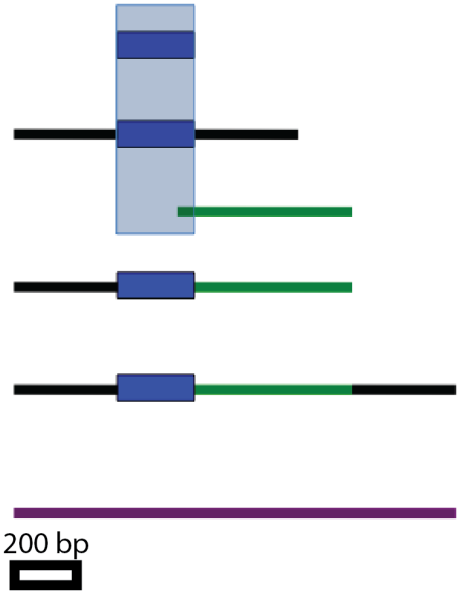
This decision was bolstered by another data set that was generated to identify potential limb-specific enhancers. Visel *et al.* (2009) used Chromatin Immunoprecipitation with sequencing (ChIP-seq) to identify non-coding regions of the mouse genome that were associated with p300, a general transcriptional activator, in dissected E11.5 limb buds(Visel et al., 2009a). Information regarding these p300-associated enhancers can also be accessed on the VEB (enhancer.lbl.gov)(Visel et al., 2007b). Enhancers involved in active transcription are associated with p300 because of the various pro-transcription functions carried out by p300: it is thought to serve as a bridge between two or more transcription factors, it can serve as a scaffold upon which numerous other transcription-related proteins are assembled, and it possesses histone acetyltransferase activity, which can modify histones and potentially other proteins as well(Chan and La Thangue, 2001; Kalkhoven, 2004). One of the 2,105 identified putative enhancers associated with p300 was a 952 bp segment that overlapped with 48 bps of the

previously described ultraconserved element, *hs280* (or, to be more specific: overlapped with the mouse ortholog of *hs280*). Therefore, two independent data sets, one based on extreme sequence conservation and one based on p300 association and active transcription in E11.5 limb buds, indicated that this genomic region was important for vertebrate limb development. To further characterize the regulatory capacity of these sequences in the mouse I treated the orthologous mouse sequences as one putative enhancer, *Mouse-280* (*M280*), which spanned the length of the ultraconserved and p300-asociated elements for a total length of 1,577 bps. Minimal additional sequence was added to this core enhancer region (Gompel et al., 2005; Pennacchio et al., 2006; Visel et al., 2007b; Cretekos et al., 2008) during the cloning process that generated the recombinant DNA used in our transgenic and gene targeting analyses (see below) so that that the total length of the interrogated element is 1,676 bps (Table 2, Fig. 1).

M1442 enhancer: Identification and validation

In addition to *M280*, I selected another enhancer from the p300 data set for further investigation. As was the case for production of transgenic mouse embryos based on ultraconserved elements, transgenic mice harboring a p300-associated element linked to *lacZ* were generated with the human orthologous p300-associated sequence. And similarly, images of these transgenic mouse embryos are available on the VEB(Visel et al., 2007b). One such group of transgenic mouse embryos, analyzed at E11.5, was made using the p300-associated element *hs1442* (*H1442-tg*). Ten of eleven embryos displayed positive *lacZ* expression in the fore- and hindlimbs (Table 2). Specifically, the consistently observed *lacZ* expression pattern was graded, with strongest intensity concentrated in the posterior half of the limb bud and then tapering off in intensity towards the middle of the bud and becoming faint to absent in the anterior half. This aspect of the expression pattern is reminiscent of *Shh* expression in the early developing limb

Figure 1. *M280* is composed of an Ultraconserved and p300-associated sequence. *M280*, the sequence used to make *lacZ* reporter mice and the sequence targeted for deletion in $M280^{tm1}$ -/- mice, is made of an Ultraconserved element (dark blue), a p300-associated element (green) and flanking sequence (black). Thus *M280* (purple) is larger than its human counterpart, *hs280*. Light blue box overlaps the Ultraconserved sequence and the region shared between the Ultraconserved and p300-associated sequences. VEB refers to the VISTA Enhancer Browser and researchers who curate it. MJN refers to author of this research.

Method of enhancer identification		Modifications / made by:	Uses
Ultraconserved (274 bp)		Flanking sequence / VEB	Transgenic mice (human sequence)
p300 - ChIP-seq		Flanking sequence / MJN	
		Flanking sequence / MJN	
		Final sequence = 1676 bp / MJN	Transgenic mice (mouse sequence); Knockout mice

bud (Lettice et al., 2003), although *Shh* expression is even more tightly restricted to the posterior edge of the limb bud. Nonetheless, the *lacZ* expression pattern driven by *hs1442* clearly overlaps with endogenous *Shh* expression. In addition to the graded posterior expression domain, *H1442-tg* embryos displayed a relatively thin expression domain that ran along the distal arc of the fore- and hindlimb buds. This portion of the pattern forms a crescent shape that encompasses the distal posterior as well as the distal anterior edges of the limb bud. The crescent is reminiscent of apical ectodermal ridge-specific expression of certain *Fgf* genes, but it is not clear from the images on the VEB that the *lacZ* expression pattern is indeed ectodermal rather than or in addition to mesenchymal.

MATERIALS AND METHODS

Generation of *M280* gene targeting vector

The *M280* gene targeting vector was engineered in four steps, using the pBluescript II KS (-) plasmid (pBS) as the vector backbone. First, a HindIII/Xho1 fragment containing a *loxP*-flanked (flox) *neo* expression cassette (from PGK-neobpA-loxP) was inserted into the HindIII and Sal1 pBS sites. Upon ligation of these two fragments the adjoining Sal1-Xho1 sites were abolished. Second, the 5'-arm of homology was inserted into the vector. Previous to insertion, the 5'-arm of homology was generated using PCR amplification with Phusion high-fidelity DNA polymerase (Finnzymes, product code: F-530). Template DNA for the amplification of the 5'-arm of homology (and the 3'-arm) was sourced from 129S6/SvEv inbred mice (Taconic). The 5'-arm of homology amplicon was designed such that the HindIII restriction endonuclease would cut an endogenous HindIII site near the 5'-end (left end) of the amplicon and also a HindIII site engineered into the amplicon via the 3'-primer (right primer). 5'-arm of homology primers are: 5'arm-L-8extra (TTTATTTGTGTGTGGTATTTGCATGTGTGC); 5'arm-R-addHindIII (TATAAGCTTCTGGTGCAAGCCTCTGGAAG). The 5'-arm of homology was inserted into the HindIII site of the vector. Third, the 3'-arm of homology was generated using the Phusion DNA polymerase. The forward primer (left end of amplicon) was designed so that a Sal1 sight was introduced into the amplicon. After amplification, a Xho1 site internal to the amplicon could be used to remove 725 bp from the 3'-end (right side) of the amplicon, priming the amplicon for ligation into the targeting vector. The 3'-arm of homology amplicon was cut with Sal1 and Xho1 and cloned into the targeting vector that had been cut with Xho1. Upon ligation of these two fragments the internal Sal1-Xho1 site was abolished. 3'-arm of homology primers are: 3armSal-Lb (ACCTGTCGACCTTGAATCTCATTCTTTTGCAGAG); 3arm-Rb

(TCTTCATGAGCTGATTAAACTCCTTGC). Fourth, a herpes simplex virus-thymidine kinase expression negative selection cassette (MCTK1; (Mansour et al., 1988)) was amplified with flanking Not1 sites using Phusion DNA polymerase and ligated into the Not1 site of the targeting vector (MCTK1 amplicon derived from pMCTK1). All components of the targeting vector were sequenced prior to combination, and the completed targeting vector was sequenced to verify that the *neo* expression cassette replaced the *M280* enhancer and that no other changes were introduced. The completed *M280* gene targeting vector was linearized using the single Kpn1 site. The 5'- and 3'-arms of homology provide 3656 and 2579 base pairs (bps) of homology, respectively (total homology = 6235 bps).

Generation of *M1442* gene targeting vector

The *M1442* gene targeting vector was engineered in four steps, using the pBluescript II KS (-) plasmid (pBS) as the vector backbone. First, the 5'-arm of homology was engineered using the Phusion DNA polymerase, with a 5'-primer containing an introduced Spe1 site and a 3'-primer containing an introduced EcoR1 site. Template DNA for the amplification of the 5'-arm of homology (and the 3'-arm) was sourced from 129S6/SvEv inbred mice. The Spe1/EcoRI 5'-arm of homology amplicon was inserted into the Xba1 and EcoR1 sites of pBS. Upon ligation of these two fragments the adjoining Spe1-Xba1 sites were abolished. 5'-arm of homology primers are: 5armL-addSpe1 (AAAACTAGTGAGCAAGGCAAGGTTTTTGG); 5armR-addEcoR1 (AAGAATTCCTCAATAATGTGGTTTCCCTTG). Second, an EcoR1/Xho1 fragment containing a *loxP*-flanked (flox) *neo* expression cassette (from PGKneobpA-lox) was inserted into the EcoR1 and Xho1 pBS sites. Third, the 3'-arm of homology was engineered using the Phusion DNA polymerase, with a 5'-primer containing an introduced Xho1 site and a 3'-primer

containing an introduced Kpn1 site. The Xho1/Kpn1 3'-arm of homology amplicon was inserted into the Xho1 and Kpn1 sites of pBS. 3'-arm of homology primers are: 3armL-addXho1 (CCACTCGAGGTTCAGCATTACTGCAAAGCAC); 3armR-addKpn1 (ATAGGTACCTATCCTGGAGCTGGAGAGATGG). Fourth, a Not1/SacII fragment containing a herpes simplex virus-thymidine kinase expression negative selection cassette (MCTK1; (Mansour et al., 1988)) was inserted into the Not1 and SacII sites of the targeting construct (The MCTK1 fragment was derived from the same amplicon described for use in the *M280* gene targeting construct). All components of the targeting vector were sequenced prior to combination, and the completed targeting vector was sequenced to verify that the *neo* expression cassette replaced the *M1442* enhancer and that no other changes were introduced. The completed *M1442* gene targeting vector was linearized using the single Kpn1 site flanking the 3'-arm of homology opposite the MCTK1 cassette. The 5'- and 3'-arms of homology provide 3047 and 3002 base pairs (bps) of homology, respectively (total homology = 6049 bps).

Generation of targeted ES cell clones and germline alleles for *M280*

The *M280* targeting vector was linearized with Kpn1 outside the 3'-arm of homology. Twenty-five µg of vector was electroporated in duplicate into PC3 cells (O'Gorman et al., 1997). Electroporated cells underwent 10 days of selection in medium that contained G418 and FIAU (Behringer et al., 1994). Three-hundred-seven G418, FIAU-resistant ES cell colonies were picked for the *M280* targeting vector and screened via mini-Southern Blot (Ramirez-Solis et al., 1992). When digested with Drd1 and hybridized with a 5'-external probe, correctly targeted ES cell clones were identified by a 15.4 kb band originating from the non-targeted locus on one sister chromosome and a second 9.6 kb band from the targeted locus on the other sister chromosome. Alternatively, the same correctly targeted ES cells could be identified by the 15.4

kb band from the non-targeted locus and an additional 5.7 kb band from the targeted locus when digested with Drd1 and hybridized with a 3'-external probe. A targeting frequency of 3.5% (11 of 307) G418, FIAU-resistant ES cell colonies screened for the 5'- and 3'-external probes was achieved. Two correctly targeted ES cell lines were injected into C57BL/6J blastocysts and transferred into 2.5-day-post-coitum pseudopregnant Swiss foster mothers. A range of chimerism from 20-90% (agouti pigmentation) was observed in the foster offspring. The targeted, PC3 ES cells are homozygous for a protamine-Cre transgene; the Cre recombinase is expressed in the male germline and results in excision of the PGK-neobpA-lox cassette and the accompanying internal Drd1 site. Excision of the PGK-neobpA-lox cassette was verified by digesting genomic DNA acquired from kidney or tail tissue from heterozygous individuals with Drd1 and hybridization to either the 5'- or 3'-external Southern Blot probes; this resulted in the visualization of 15.4 and 13.7 kb bands. Deletion of the *M280* enhancer was verified by digesting genomic DNA with Drd1 and hybridization with a Southern Blot probe complimentary to the sequence of the enhancer itself. Thus in homozygous individuals no 13.7 kb band was observed. Male chimeras were bred to B6 females, and a subset of agouti offspring were genotyped using the Southern Blot procedure outlined above to confirm germline transmission of the targeted *M280* allele. Confirmed transmitting male chimeras were then bred to 129S6/SvEv females to maintain the lines on a 129S6/SvEv inbred background. The F1 progeny from the crosses, and all other crosses, were genotyped by PCR using DNA collected from tail biopsies. The PCR reaction consisted of three primers that could distinguish the three genotypes, non-targeted, heterozygous targeted, and homozygous targeted: 280-PreEnh-Lb (GGTTCAATCAGCCTTTCATTCCAG); 280-wtEnh-R6 (AGGCTTCCAGGGCTGATAACAAG); 280-PostEnh-Rd (AGGCAGCTCTGTCTCTGCAAAAG). Primer set 280-PreEnh-Lb + 280-wtEnh-R6 provides a 405 bp amplicon that signifies the non-targeted allele. Primer set 280-PreEnh-Lb + 280-

PostEnh-Rd provides a 248 bp amplicon when amplified across the deleted *M280* allele; on the non-targeted allele the 280-PreEnh-Lb + 280-PostEnh-Rd primer set does not produce an amplicon when using Choice *taq* DNA polymerase (Denville Scientific Inc.). Note that because the PC3 cells used for electroporation of the targeting vector are homozygous for the protamine Cre, all F1 progeny from the male chimeras X 129S6/SvEv females will be heterozygous for the Cre; thus F1 progeny must be crossed to additional 129S6/SvEv individuals and genotyped for the presence/absence of the Cre transgene.

Generation of targeted ES cell clones and germline alleles for *M1442*

The *M1442* targeting vector was linearized with Kpn1 outside the 3'-arm of homology. Twenty-five μg of vector was electroporated in duplicate into PC3 cells (O'Gorman et al., 1997). Electroporated cells underwent 10 days of selection in medium that contained G418 and FIAU (Behringer et al., 1994). Two-hundred-twenty-two G418, FIAU-resistant ES cell colonies were picked for the *M1442* targeting vector and screened via mini-Southern Blot (Ramirez-Solis et al., 1992). When digested with Xba1 and hybridized with a 5'-external probe, correctly targeted ES cell clones were identified by a 14.4 kb band originating from the non-targeted locus on one sister chromatid and a second 6.6 kb band from the targeted locus on the other sister chromatid. Alternatively, the same correctly targeted ES cells could be identified by the 14.4 kb band from the non-targeted locus and an additional 6.8 kb band from the targeted locus when digested with Xba1 and hybridized with a 3'-external probe. A targeting frequency of 2.7% (6 of 222) G418, FIAU-resistant ES cell colonies screened for the 5'- and 3'-external probes was achieved. Two correctly targeted ES cell lines were injected into C57BL/6J blastocysts and transferred into 2.5-day-post-coitum pseudopregnant Swiss foster mothers. A range of chimerism from 50-80% (agouti pigmentation) was observed in the foster offspring. The targeted, PC3 ES cells are

homozygous for a protamine-Cre transgene; the Cre recombinase is expressed in the male germline and results in excision of the PGK-neobpA-lox cassette and the accompanying internal Xba1 site. Excision of the PGK-neobpA-lox cassette was verified by digesting genomic DNA acquired from kidney or tail tissue from heterozygous individuals with Xba1 and hybridization to either the 5'- or 3'-external Southern Blot probes; this resulted in the visualization of 14.4 and 11.7 kb bands. Deletion of the *M1442* enhancer was verified by digesting genomic DNA with Xba1 and hybridization with a Southern Blot probe complimentary to the sequence of the enhancer itself. Thus in homozygous individuals no 11.7 kb band was observed. Male chimeras were bred to C57BL/6J females, and a subset of agouti offspring were genotyped using the Southern Blot procedure outlined above to confirm germline transmission of the targeted *V442* allele. Confirmed transmitting male chimeras were then bred to 129S6/SvEv females to maintain the lines on a 129S6/SvEv inbred background. The F1 progeny from the crosses, and all other crosses, were genotyped by PCR using DNA collected from tail biopsies. The PCR reaction consisted of three primers that could distinguish the three genotypes, non-targeted, heterozygous targeted, and homozygous targeted: 1442-PreEnh-L2 (AAAGCCCAAAGGTTTTTCTCAAGG); 1442-wtEnh-R1 (CAAACATCTCCTACAAGCACTCACG); 1442-PostEnh-R2 (GTGAACTCGGCAGGTAACAGACAG). Primer set 1442-PreEnh-L2 + 1442-wtEnh-R1 provides a 496 bp amplicon that signifies the non-targeted allele. Primer set 1442-PreEnh-L2 + 1442-PostEnh-R2 provides a 264 bp amplicon when amplified across the deleted *M280* allele; on the non-targeted allele the 1442-PreEnh-L2 + 1442-PostEnh-R2 primer set does not produce an amplicon when using Choice *taq* DNA polymerase (Denville Scientific Inc.). Note that because the PC3 cells used for electroporation of the targeting vector are homozygous for the protamine Cre, all F1 progeny from the male chimeras X 129S6/SvEv females will be

heterozygous for the Cre; thus F1 progeny must be crossed to additional 129S6/SvEv individuals and genotyped for the presence/absence of the Cre transgene.

Generation of *M280* and *M1442* transgene constructs and associated transgenic mice

Transgenic mice were produced with two general types of constructs: 1) constructs harboring either *M280* or *M1442*, the minimal heat shock protein 68 promoter (HSP68), and the *E. coli β-galactosidase* (*lacZ*) reporter gene [phspPTlacZpA; (Kothary et al., 1989)]; or 2) constructs with the same configuration as in "1" but having been built using the Invitrogen Gateway Technology. Constructs in category 1 were used as templates to amplify the enhancer-fragments used in category 2. Enhancer fragments were amplified from genomic DNA from 129S6/SvEv mice using Phusion High-fidelity DNA polymerase (Finnzymes, F-530) with primers designed to have flanking HindIII sites. The 2723 bp *M1442* amplicon was inserted into the phspPTlacZpA HindIII site. *M1442* transgene primers are: 1442-tg-L (TTTAAGCTTAGGGAAACCACATTATTGAG), 1442-tg-R (GTTAAGCTTGCTTTGCAGTAATGCTGAAC). Excision of the transgene away from the plasmid backbone was achieved by cutting with Sal1. The 1676 bp *M280* amplicon was blunt-end cloned into the phspPTlacZpA Sma1 sites because after amplification it was noticed that there was an internal HindIII site in the *M280* vector. *M280* transgene primers are: 280-tg-L (TGGAAGCTTCTTCCAGAGGCTTGCACCAG), 280-tg-R (TCTAAGCTTTTGCTCTCTGTCCCCTGTCC). Excision of the transgene away from the plasmid backbone was achieved by cutting with Sal1.

To generate the Invitrogen Gateway-based transgenes, *M280* and *M1442* PCR-generated amplicons were obtained from the transgene vectors described above. Left and right primers (5' and 3') for both *M280* and *M1442* amplicons were designed with flanking Xho1 sites for easy

insertion into the pENTR 3C (Invitrogen, A10560) Xho1 site. *M280* Gateway transgene-production primers are: 280-Xho-L (TGCCTCGAGCCTGCAGTCTCTCCCC), 280-Xho-R (GGTCTCGAGGTCAGTATCATGACC). *M1442* Gateway transgene-production primers are: 1442-Xho-L (GTTCTCGAGACTTTCATTTAGGAGG), 1442-Xho-R (CTGCTCGAGCTACACAAACATGA). The pENTR 3C multi-cloning site is flanked by the attL1 and attL2 recombination sites. Upon insertion of either enhancer fragment the 5' pENTR 3C Sal1 site (nearest to attL1) was destroyed. Using the Invitrogen Clonase II Plus Enzyme Mix (12538-120) the enhancers were then transferred from the pENTR 3C vector to a plasmid containing Gateway cloning sites (attL1/attL2) upstream of HSP68 and *lacZ* (kindly provided by Nadav Ahituv, University of California San Francisco; based on initial vector in (Kothary et al., 1989)). The Sal1 site just 5' of the attL2 recombination site in the pGW-HSP-LZ vector was removed via the LR Clonase recombination event, therefore both the *M280* and *M1442* Gateway-based transgenes were excised away from the pGW-HSP-LZ backbone via two Sal1 sites that flank the entire transgene. Each transgene was sequenced before excision and preparation for pronuclear injection into mouse zygotes. Transgenic mice were generated with both the *M280* and *M1442* enhancers in either the phspPTlacZpA or pGW-HSP-LZ vector via pronuclear microinjection (Brinster et al., 1985) into zygotes obtained from either FVB/N intercrosses (used for the stable transgenic lines characterized herein) or [C57BL/6 X SJL]F_1 hybrid intercrosses. Generation zero (G_0), E11.5 embryos harboring the phspPTlacZpA-based transgenes were collected from foster mothers and analyzed for *lacZ* expression by X-gal staining. Stable transgenic lines were generated using the pGW-HSP-LZ-based vectors. Six of twenty G_0 mice screened were positive by PCR for the *M280*-pGW-HSP-LZ transgene; of these, 3 founders (50%) expressed *lacZ* (as assayed by X-gal staining) in developing limbs. *M280*-pGW-HSP-LZ transgene PCR genotyping primers are: 280GWtg-L (CTCTTGGGCCTTGGGGATAGTAG), 280/1442GWtg-R2

(AGTGCTGCCTCTGACCTCATGG). Eight of twenty-nine G_0 mice screened were positive by PCR for the *M1442*-pGW-HSP-LZ transgene; of these, 2 founders (25%) expressed *lacZ* (as assayed by X-gal staining) in developing limbs. *M1442*-pGW-HSP-LZ transgene PCR genotyping primers are: 1442GWtg-L2 (AATGTCCTCTGACCTCCACATGC), 280/1442GWtg-R2 (AGTGCTGCCTCTGACCTCATGG).

Quantitative real-time PCR analysis

M280$^{tm1/tm1}$, M1442$^{tm1/tm1}$, or 129S6/SvEv (wild type) mice were intercrossed and E11.5 embryos were collected from sacrificed females. Forelimb and hindlimb autopods (hands and feet) were dissected away from three embryos of the same genotype and pooled together (1 sample = 12 autopods; 6 forelimbs and 6 hindlimbs). Pooled autopods were moved into 500 µl Ambion TRIzol Reagent (15596) and homogenized using sterile 1mL syringes and 27G x ½ (0.4mm x 13mm) needles. Total RNA was obtained following the manufacture's guidelines (Ambion TRIzol Reagent, 15596) and RNA quality and concentration was assessed on a NanoDrop Spectrophotometer (ND-1000; 260/280 nm ≥ 1.8; 260/230 nm ≥ 1.9). Five hundred ng of total RNA from each sample was used to make cDNA with the Invitrogen SuperScript III First-Strand Synthesis System for RT-PCR (18080). Quantitative RT-PCR (qPCR) was carried out in duplicate or triplicate on oligo-dT-primed cDNA (n=8 biological replicates per genotype) on an ABI 7500 Real-Time System (Applied Biosystems). Initially, numerous endogenous controls were tested: *Gapdh, Mouse-actin, Mouse-18S ribosomal RNA, Mll* and *Rbbp4*. Results calculated with reference to different endogenous control genes did not always agree. For example, one endogenous control may provide non-significant transcriptional change for one gene of interest (GOI), while another endogenous control would provide a significant up- or down-regulation. I decided to dispense with *Gapdh* because it provided very inconsistent results

over several trials. I also dispensed with *Mouse-actin* because I reasoned that in a developing tissue, such as a E11.5 limb bud, that cell movement and growth may lead to fluctuating levels of cell cortex proteins such as actin. Ultimately I settled on using *Rbbp4* because it provided consistent results in two or three trials per GOI. And most importantly, *Rbbp4* was originally selected because of its consistent transcriptional levels across numerous amniocyte stem cell lines (Maguire et al., 2013) (special thanks to Colin T. Maguire, University of Utah, for directing me to his publication and for suggestions concerning qPCR internal controls). The comparative C_T method was employed to determine the relative expression of numerous GOI after normalization to the *Rbbp4* endogenous reference and calibration to wild-type GOI expression levels (Applied Biosystems, Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR). Primer sequences for genes assessed are in Table 3 below. The primer set for *Prx1* was obtained from Cretekos, C. *et al.* (2008). All other primer sets were designed using Oligo Architect (http://www.sigmaaldrich.com/life-science/custom-oligos/dna-probes/product-lines/probe-design-services.html) or the Roche Universal ProbeLibrary Assay Design Center (https://www.roche-applied-science.com/webapp/wcs/stores/servlet CategoryDisplay?catalogId =10001&tab=Assay+Design+Center&identifier=Universal+Probe+Library&langId=-1). All primer pairs were designed to flank at least one intron (if possible); additionally, most primer sets were designed such that at least one primer annealed to an exon-exon junction.

Whole Mount In Situ Hybridization (WMISH)

WMISH was carried out using standard protocols (Nagy et al., 2003). Embryonic day 10.5 and 11.5 embryos were collected several weeks or months before the WMISH experiment from mutant or unaltered 129S6/SvEv intercrosses. Embryos were fixed in 4% paraformaldehyde

Table 3. Primer sets for Quantitative Real-time Polymerase Chain Reaction. All sequences read from 5' to 3'.

Gene Name	Forward Primer	Reverse Primer
Atpb1	TCAGGTTCCTGGAAAAGTACAAG	TTGGGTTCACTGGGAACATT
BC055324	GTAAGAACAGCGAATGTG	TGAATGAATAACTCCTACCA
Dmap1	TTAAGTCGGCAGGTGTCA	GCTCCACACCAAGTTCCA
Eri3	GATGAGTGGATGGCGAAGG	GCACTGTCCTGGAAGCAT
Gorab	CAAGAACAGCGACTAATGGA	CCTGAGGATGCCAATGTC
Kifap3	CAGTGATGAAGAGGAGGAGTT	GGCACCTGGCTTTAGTTTG
Nme7	GCTGGTGTTGATTGACTATGG	TCCGCTGCTTCTTTCCTT
Prdx1	CCGCTCTGTGGATGAGATTAT	GTATCACTGCCAGGTTTCCAG
*Prx1**	GATCTCGCACGTCGGGTGAACC	GCCAGCATGGCTCGCTCATTC
Rbbp4	CTTAGTGCTTCAGATGACCATACC	GCATCCACCACCTTTCCTT
Rnf220	CAGCGACATTGAGAAAATC	CAGCCAACACTCTTCACA
Scyl3	GAATATCCAGTCAGTAAGAGA	AACCTGCTCATTGAAGAT
Tmem53	CCAAGTATAGTGCTATCTACC	CCGCTCAATCTCATAGTC

*Primer set obtained from Cretekos, *et al.* (2008).

overnight, rinsed three times in PBS-Tween-20 (PBT), and walked through a series of methanol:PBT washes until dehydration in 100% methanol. Embryos were then stored at -20 ° C. Both sense and antisense *Rnf220* riboprobes were amplified from cDNA collected from E12.5 129S6/SvEv whole embryos and then cloned in pBluescript II KS (-). *Rnf220* riboprobe primers were obtained from Mouse Genome Informatics website (www.informatics.jax.org; primer set ID: MTF1124). The *Shh* riboprobe was obtained from Andrew McMahon's laboratory (Echelard et al., 1993). I designed the primer sets for all other riboprobes used in this study (*Kifap3, Dmap1, Nme7*) using Oligo Architect (http://www.sigmaaldrich.com/life-science/custom-oligos/dna-probes/product-lines/probe-design-services.html). Primer sets are provided below in Table 4. As with production of the *Rnf220* riboprobe, the probes for *Kifap3, Dmap1, Nme7* were amplified from cDNA collected from E12.5 129S6/SvEv whole embryos and cloned into pBluescript II KS (-), except for the *Rnf220* riboprobe, which was cloned into pGem-T Easy (Promega).

Forelimb autopod bone measurements

Embryonic day 18.5 forelimbs stained with alcian blue and alizarin red from $M280^{tm1/tm1}$, $M1442^{tm1/tm1}$, or 129S6/SvEv (wild type) mice were cut so as to separate the forearm, which includes the autopod/stylopod (radius + ulna + hand), from the hind arm. Genotyped limbs were obtained from heterozygous intercrosses between either $M280^{tm1/+}$ or $M1442^{tm1/+}$ mice, or from homozygous intercrosses between either $M280^{tm1/tm1}$, $M1442^{tm1/tm1}$, or 129S6/SvEv (wild type) mice. The forearms were then placed, palm-side down, into the 13 mm diameter, 1 mm depth well of a 25 x 25 mm silicon isolator (Electron Microscopy Sciences, cat. # 70336-02) that was fitted onto a 25 x 75 x 1 mm microscope slide The small well in the middle of the silicon isolator was filled with 1:1 95 % ethanol:glycerol; the well with forearm was capped with a 22

Table 4. Primer sets for Whole Mount *In Situ* Hybridization. All sequences read from 5'-3'.

Targeted transcript/ Sequence cloned into vector	**Adaptor Restriction Enzyme Sites Engineered into Primers: Forward/ Reverse**	**Forward Primer**	**Reverse Primer**
Dmap1	SacII (Rev.)	CGTAGTTATTCAC GATCGGTATGACCAC	AGTCCGCGGATTC AGACACTGTTGGCTCAG
Kifap3	XbaI/XbaI	TGTCTAGAAAAGA CACGAGCTTTGGCAAG	AATCTAGACTTA AAGCGGTCATCCATGC
Nme7	XhoI/PstI	GCACTCGAGAATC AGAGCGAGAGATTCG	CTCTGCAGCAGATG CAAAAGTATCAGG
*Rnf220**	N/A	ACGGAGGAAGC AAGATGAAGGG	CCGAGATAGCTG CCGTTCAAGC
*Shh***	N/A	N/A	N/A

*Obtained from www.informatics.jax.org; primer set ID: MTF1124

**Obtained from A. McMahon Lab (Echelard et al., 1993)

by 22 mm slide cover. Light pressure was applied to the slide cover to ensure that the forelimbs from each individual were oriented uniformly. Images of limbs were taken on a Leica M205 Fluorescence Microscope with fluorescent filters removed (brightfield) at the same magnification and settings. Because the *lacZ* expression patterns driven by both *M280* and *M1442* did not intersect with the developing stylopod (radius and ulna) we expected the morphology of the radii to be unaffected by the enhancer deletions in our two lines of targeted mice. Upon measuring the mineralized portion (stained red by alizarin red) of E18.5 radii we found there to be no statistical difference in the length of radii between genotypes when corrected for body weight. Therefore, all of our subsequent measurements of autopod bone elements were normalized to an individual's radius length. Bone lengths were measured in both right and left autopods in Photoshop using arbitrary units with the ruler tool. Measurements for right and left bone elements were averaged and used in the final calculations.

In accordance with the expression pattern in our *M280*-pGW-HSP-LZ transgenic mice, bone elements of forearm digit III were measured in E18.5 $M280^{tm1/tm1}$ embryos and compared to wild-type controls: third metacarpal, phalanges 1-3. Digit III bone elements (mineralized as well an un-mineralized portions) were measured individually and then length measurements were combined for a total length that was normalized to radius length.

In accordance with the expression pattern in our *M1442*-pGW-HSP-LZ transgenic mice, the mineralized as well an un-mineralized portions of forearm digits I and IV were measured in E18.5 $V442^{tm1/tm1}$ embryos and compared wild-type controls: phalanges 1-2 for digit I; fifth metacarpal and phalanges 1-3 for digit IV. Bone elements for each digit were measured individually and then length measurements were combined for a total length that was normalized to radius length.

Measurements of E15.5 Crown-rump (CR), mandible, ulna and humerus lengths

Images of whole E15.5 embryos were taken on a Leica M205 Fluorescence Microscope with fluorescent filters removed (brightfield) at the same magnification and settings. In Adobe Photoshop the length of the desired cartilage element (e.g. CR-length or mandible length) was measured using the ruler tool and lengths are reported as "arbitrary length units."

Limb-only bone and cartilage stain

Collected E18.5 embryos were placed in a 100 mm petri dish, sealed with paraffin, and left at -20°C overnight. (The E18.5 time point was chosen to avoid weight and size differences that may have resulted from chance differences in nursing opportunities). The next day embryos were put into water that had cooled for one minute after boiling. Embryos were left in the water for 15-30 seconds, until skin turned a milky white (blanched). After briefly patting the embryos dry, the skin was completely removed from the fore- and hindlimbs and then the limbs were dissected off the rest of the embryo: Scapula to digits for forelimb, and tibia to digits for hindlimb. Limbs were placed in 95% ethanol overnight without rocking to avoid damage to the digits. Limbs were then moved to 25 mm petri dishes filled with filtered alcian blue stain (20% acetic acid, 0.05% alcian blue in 95% ethanol; filtered with Whatman paper) and left overnight, or for approximately 12 hours. (To keep the limbs from being jostled too much during solution changes they were placed inside open tissue cassettes that sat inside the petri dishes). After alcian blue staining the limbs were washed several times in 95% ethanol over five hours. Cover the limbs in 1.5% potassium hydroxide (KOH) for one hour, then change to 0.5% KOH and leave overnight. Early the next morning change the solution to alizarin red stain (0.5 % KOH, 0.015% alizarin red in water), completely covering the limbs, and before leaving for the day

change the solution to 0.5% potassium hydroxide/20% glycerol. Leave limbs in 0.5% potassium hydroxide/20% glycerol for one full day. When samples are appropriately cleared (~24 hours), place them in glycerol:95% ethanol (1:1) for documentation and storage.

Embryonic cartilage staining; limb alignment; assay for transient developmental defects

Embryonic day 13.5-15.5 embryos were stained with alcian blue to visualize cartilage as described elsewhere (Nagy et al., 2003), with the following changes: E13.5 embryos stained in alcian blue for 2-3 hours, E14.5 embryos stained for 4-5 hours, and E15.5 embryos for 6.5-7 hours. Left limbs are dissected away from stage-matched embryos from and imaged with a Leica M205 Fluorescence Microscope with fluorescent filters removed (brightfield) at the same magnification and settings. In Adobe Photoshop limbs are aligned together in one file, with the proximal tip of the ulna as the reference point. Embryonic limbs are then traced in Adobe Illustrator and overlaid onto one another. Because of minor but common differences in developmental maturation between litters, mutant and control embryos were compared only if from the same litter. In this manner I could assess whether mutant limb elements were larger or smaller or aberrantly shaped relative to controls. Similar to the analysis performed with E18.5 digit lengths, I assessed developmental progression of autopod elements that would likely be affected by the absence of the deleted enhancers based on the expression patterns in my driven by each enhancer in our stable transgenic lines. Mutant and control autopods were scored as equal or non-aberrant if half the cartilage elements under consideration, plus one, were of equal size and shape when their outlines were overlaid on one another. Equality was also scored if there was a combination of equal, large and small mutant elements within a single pairwise comparison, with the count of any category never exceeding the majority of elements under consideration. In contrast, mutants were scored as aberrant and smaller than controls if one half

of the elements under consideration, plus one, were smaller in size and shape that control elements. Aberrantly large mutant autopods were scored in the same manner as aberrantly small autopods.

X-gal and Salmon-gal detection of transgene activity

Either X-gal (Goldbio.com; Cat # X4281C) or Salmon-gal (LabScientific, Inc; Cat# X668) was used as a substrate to detect *lacZ* expression in my transgenic mouse lines following established procedures (Nagy et al., 2003; Sundararajan et al., 2012). Both reactions were carried out either the same day or the day after collection of embryos. Salmon-gal was used only for the weakly expressing *M280*GWtg lines (see Results chapter). When using the Salmon-gal protocol (Sundararajan et al., 2012) I left embryos in Salmon-gal Staining Solution for 30 minutes. Embryos were fixed before the staining reaction in 4% paraformaldehyde, but were not post-fixed. After X-gal or Salmon-gal staining embryos were washed in PBS, imaged (in PBS) if necessary, and stored in 70% ethanol.

Nuclear Fast Red, and Hematoxylin and Eosin staining of embryonic limb sections

To assess the extent of *lacZ* expression in my transgenic lines, embryos from E11.5-E15.5 were embedded in Paraffin wax and sectioned at between 6 and 12 microns. Sectioned material was counterstained with Nuclear Fast Red (Sigma-N3020) for approximately 3-5 minutes. Embryonic day 18.5 limbs were collected for histological sections using hematoxylin and eosin (H&E) stains. Prior to embedding into Paraffin wax, E18.5 limbs were fixed in 10% Formalin overnight, washed three times in PBS for at least 10 minutes, and rinsed in 0.5M EDTA for 48

hours to remove calcium deposition of the bone. Sections of E18.5 limbs ranged from 6-10 microns.

Method for identifying potential shadow enhancers near *M280*

The sequence for *M280* was entered into the ECR Browser (http://ecrbrowser.dcode.org) using the appropriate genomic coordinates (mm9). It was noticed that extending from mouse to *Xenopus laevis* (frog) there are two conserved peaks within this regulatory region [520 (large) and 148 (small) bps, respectively]. Both the large and small sequence blocks derived from the frog genome were selected by clicking; within the new dialogue box the TFBS option in the upper right was selected. Once taken to the webpage "http://rvista.dcode.org/cgi-bin/rFromXB.cgi" the default parameters were left as they were (i.e. vertebrate, etc.). All TF binding sites (TFBS) were selected and submitted (i.e. SELECT ALL option). A list of 184 conserved putative TFBS was returned. This list was narrowed down to 14 TFBS by looking for an association between the name of the TF and the key words "limb, development, mouse" on both Pubmed and GoogleScholar (Appendix 1). These 14 TFBS then allowed me to interrogate other forms of putative regulatory sequences around *M280* (p300-bound sites and highly conserved human-mouse-rat sequences); the prediction was that shadow enhancers for *M280* would harbor similar clusters of these 14 TFBS.

The list of putative *M280* shadow enhancers ± 1 Mb from *M280* was generated from two previously published datasets: A) limb-specific p300 binding sites (N=6), see Supplemental Material in Visel, *et al.,* 2009); and B) regions of highly constrained conservation between humans, rats, and mice (N=12; although not ultraconserved), see Supplemental Material in Visel, *et al.,* 2008. FASTA files for each putative shadow enhancer were genetated; sequences used were from mouse genomome build mm9 (UCSC Genome Browser). Mouse genome

coordinates for each putative shadow enhancer were then submitted to the ECR browser; which generated a visual readout of multi-species alighnements. Aligned sequences corresponding to hg19 (human genome; (which had been aligned to the mm9 base sequence on the ECR Browser) were then clicked. In the top right of the new dialogue box the tab labeled "conserved transcription factor binding sites" was selected. This opens up a new window (http://rvista.dcode.org/cgi-bin/rFromXB.cgi). In the new rVista widow I clicked Submit. In the next window I selected the following TFBS based on their association with limb development as previously described: AP2ALPHA, CHX10, CLOX, E2F1_Q3, E2F1DP1RB, E2F4DP1, E2F_Q3, ETS_Q4, ETS_Q6, FOXM1, GATA2, HOXA4, NRF2, PEA3. I then clicked Submit and then, in turn, Check It. To collect the list of conserved TFBS I clicked on the tab labeled "Conserved" under the "Summary" option. The output file generated permitted me to visualize the conserved TFBS between mm9 and hg19. In the readout the first sequence is the mouse sequence, the second is the human sequence. The best candidate shadow enhancers for *M280* were selected based on their possession of clusters of TFBS originally identified in the large and small block of conservation descrived in above.

RESULTS

As was briefly explained above, the enhancers I selected to characterize were based largely on images of previously generated transgenic mice (Table 2). These images indicated at least partially, through the presence of a well-defined *lacZ* expression pattern, the regulatory capabilities of the enhancers. Importantly, the imaged transgenic mice were made using human genomic sequences (Pennacchio et al., 2006; Visel et al., 2007b; Visel et al., 2009a) and were scored for limb-specific (and other tissues) *lacZ* expression at a single time point, E11.5. Therefore, it was necessary to assess whether the orthologous mouse enhancer behaved in a similar manner. To do this I generated a series of transgenic mice that harbored transgenes wherein the reporter gene, *lacZ*, was driven by either *M280* or *M1442*, and a minimal promoter, *HSP68* (Kothary et al., 1989) (Table 5).

Gross analysis of *M280* transgenic mice

Initially, *M280* was cloned upstream of the minimal promoter, *HSP68*, and the reporter gene, *lacZ*, using traditional cloning methods. The linearized and purified transgene was then injected into F_2 zygotes from (C57BL/6 X SJL)F_1 intercrosses and transferred to pseudopregnant Swiss females via oviduct transfer. As a first pass to determine whether *M280* behaved similarly to *hs280,* I collected E11.5 embryos (G_0 embryos) from foster mothers and scored them for the presence of the *M280* transgene via *lacZ* staining (Fig. 2B).

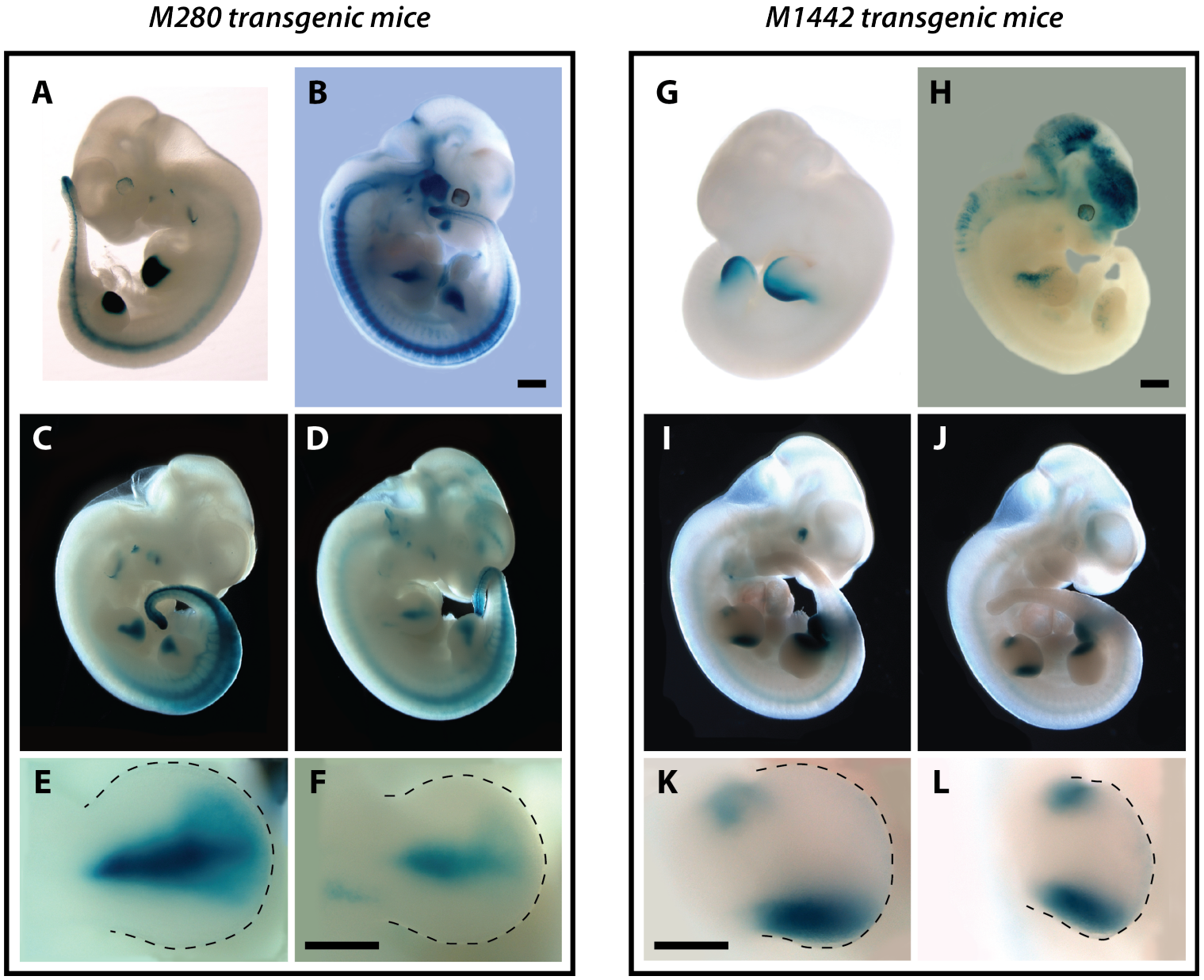
Embryos with a positive *lacZ* signal possessed an expression pattern comparable to that seen in *hs280*-transgenic mice from the VEB (Fig. 2A, B). The expression pattern was triangular or wedge-shaped in appearance and was centered in the middle of the autopod portion of the limb bud. The autopod-centered expression pattern in the *M280*-transgenic mice differed from

Table 5. Summary of *M280* and *M1442* transgenic mice.

Summary of transgenes constructed with *M280*				
Vector backbone used	**Zygote strain for pronuclear injection**	**# *lacZ* positive / genotyped positive for transgene**	**Embryos collected at E11.5 (G_0) or stable line produced**	**Expression pattern in limb buds at E11.5 similar to that reported on the VEB (see Table 4 and Figs. 2)^**
phspPTlacZpA*	[C57BL/6 X SJL]F_1 Intercross	2/2	G_0	1 embryo similar; 1 embryo dissimilar (expression throughout limb bud)
pGW-HSP-LZ**	FVB/N	7/8	G_0	Similar
pGW-HSP-LZ**	FVB/N	3/5	Stable	Similar
Summary of transgenes constructed with *M1442*				
phspPTlacZpA*	[C57BL/6 X SJL]F_1 Intercross	1/1	G_0	Dissimilar (no expression in posterior limb bud)
pGW-HSP-LZ**	FVB/N	4/4	G_0	Similar[Ŧ]
pGW-HSP-LZ**	FVB/N	2/8	Stable	Similar[Ŧ]

*See Kothary *et al.,* 1989
** See Visel *et al.,* 2007b; Visel *et al.,* 2008
^See also Figs. 4, 5, 7 and 9
[Ŧ] In addition to the predicted posterior expression pattern, forelimbs also showed *lacZ* expression in the anterior margins of the autopod (see text).

Figure 2. G_0 *M280* and *M1442* Transgenic mice collected at E11.5. (A-F) Related to *M280* enhancer. (G-L) Related to *M1442* enhancer. (A) Image of *hs280* transgenic embryo taken from the VEB (enhancer.lbl.gov; Visel *et al.* 2007b). (B) Transgenic embryo that harbors transgene composed of *M280*, *HSP68* and *lacZ.* (C-F) Representative transgenic embryos that harbor transgene composed of *M280* flanked by Gateway recombination sites (see text), *HSP68* and *lacZ.* (E, F) Close-up of right limb buds. (G) Image of *hs1442* transgenic embryo taken from the VEB (enhancer.lbl.gov; Visel *et al.* 2007b). (H) Transgenic embryo that harbors transgene composed of *M1442*, *HSP68* and *lacZ.* (I-L) Representative transgenic embryos that harbor transgene composed of *M1442* flanked by Gateway recombination sites (see text), *HSP68* and *lacZ.* (K, L)) Close-up of right limb buds. Scale bars in (B) and (H) are 0.5 mm. Scale bars in (F) and (K) are 0.25 mm.


M280 transgenic mice
M1442 transgenic mice
A
B
C
D
E
F
G
H
I
J
K
L

that in their *hs280* counterparts in that the intensity of the *lacZ* signal diminished towards the distal edge of the autopod, whereas the four *hs280*-transgenic embryos on the VEB showed strong *lacZ* expression extending from the center of the autopod to its distal borders. Nonetheless, our G_0 *M280*-transgenic mice confirmed that *M280* possesses similar regulatory capabilities as its orthologous human sequence. The differences in limb bud *lacZ* expression could be due to sequence diversity throughout the entire length of the orthologous sequences or to the site of transgene insertion, which could influence the activity of the enhancer.

My *M280* G_0 E11.5 transgenic embryos also manifested *lacZ* expression in locations other than the limb, including the neural tube, the isthmus between mid- and forebrain, the oral groove region of the first pharyngeal arch and the tail bud. Some of the extra-limb patterns of expression are also seen in the *hs280*-transgenic mice on the VEB website, including the neural tube, tail bud and regions of the first-through-third pharyngeal arches (Fig. 2).

After generating the *M280* transgene using traditional cloning methods I decided to make another construct that was a more faithful mimic of that used to generate the transgenic mice on the VEB. The human-sequence-based transgenic mice I was comparing our transgenic embryos to contained transgenes composed of the minimal promoter, *HSP68*, and *lacZ* as well, but the enhancers were inserted into the transgene vector using the Invitrogen Gateway Technology cloning system rather than with traditional cut-and-paste cloning techniques (Ko, 2001). In the Gateway system, fragments of interest are transferred from one plasmid to another via site-specific recombination events mediated by agents based on the activity of the bacteriophage lambda. Although the recombination event that transfers, say, an enhancer sequence, into an expression vector containing *HSP68* and *lacZ*, does not require the use of restriction enzymes it does leave behind short sequences from the recombination sites (various types of *att* sites) that flank the transferred sequence – the enhancer in this case. Therefore, to recapitulate the transgenic environment used in the embryos on the VEB as faithfully as

possible I decided to re-construct the *M280*-transgene using the Gateway system. This enabled me to have more confidence that any similarities or differences observed in the expression patterns of the *hs280*- and *M280*-transgenic mice were due to the species-specific enhancers and not to the flanking transgene environments.

To make a new transgene utilizing Invitrogen Gateway Technology (*M280*-GWtg) I subcloned *M280* into the Gateway vector that would then permit me to move the enhancer to any other Gateway vector of my choice (pENTR 3C, Invitrogen, A10560). The *M280* enhancer was then moved to the same vector used to generate the human-based transgenic mice observed on the VEB (Pennacchio et al., 2006; Visel et al., 2007b), which contained a Gateway cloning site, the *HSP68* minimal promoter and the *lacZ* reporter gene.

As was done for the original *M280*-transgenic embryos, *M280*-GWtg embryos were initially collected at E11.5 (G_0) and scored for their *lacZ* expression patterns in limb buds to verify that the *M280* regulated a well-defined, limb-specific expression pattern in the context of the Gateway transgene similar to that observed in *hs280*-transgenic embryos. Positive *lacZ* expression was scored in eight embryos from this experiment, seven of which exhibited a reproducible expression pattern in both the fore- and hindlimb autopods (Fig. 2C-F). As was the case for previously described expression patterns driven either by *M280* or *hs280*, these seven transgenic embryos' *lacZ* expression was triangular or wedge-shaped, with the base of the triangular expression pattern running along the distal margins of the autopod (Fig. 2E, F). In strongly expressing individuals, the *lacZ* expression pattern was most intense in the middle of the autopod; additionally, the intensity of the signal was maintained up to the distal border of the autopod in a region that would correspond to the location of digit III. Moderately intense *lacZ* expression was also detected in the regions that would likely give rise to digits II and IV. In weakly expressing individuals the *lacZ* expression pattern was mostly confined to the middle of

the autopod, and in some cases extended towards the distal autopod margin, though faintly, into the region of future digit I.

Seven of the eight positively expressing embryos had *lacZ* expression in the tail bud and in the pharyngeal arches. Particularly consistent, is the expression in the proximal margins of the oral groove of the first pharyngeal arch. Half of the embryos displayed neural tube expression that extended from at least the hindbrain through to the tail, and five of the eight embryos appeared to have *lacZ* expression that outlined the somites in the posterior half of the embryo. In some of these cases it appeared that the nerves emanating from the developing neural tube (spinal cord) were also *lacZ* positive. Therefore, though *M280* directs gene expression in a specific, delimited manner in E11.5 limb buds it also possesses the capacity to consistently direct expression in at least the E11.5 neural tube, tail bud and first pharyngeal arch (Fig. 2C, D).

From these observations I concluded that *M280* can consistently drive gene expression in the middle of the autopod at E11.5 in regions that approximately correspond to the future locations of digits I-III and perhaps to regions that will give rise to the distal portions of wrist and ankle elements. Consistent with the intent of this project, I determined to focus my assessment of *M280* as it relates to limb development, leaving close attention to the extra-limb expression patterns for a later time. With this information in hand I decided to generate stable *M280*-GWtg lines to assess the regulatory capacity of *M280*, specifically in the limb, throughout development.

I designed a primer set to specifically detect the presence of the *M280*-GWtg. Of five *M280*-GWtg positive mice, three produced embryos that had detectable *lacZ* expression in E11.5 limbs. Two of these three stable lines expressed *lacZ* relatively weakly. Therefore, only two lines were used to characterize *M280*'s regulatory capacity throughout development, one of the weak lines (*M280*-GWtg-Line 1, or just *M280*L1) and the remaining, relatively strongly

expressing line (*M280*-GWtg-Line 2, or just *M280*L2). Because neither of our stably expressing lines showed a *lacZ* expression pattern as intense as those observed in our G_0 embryos, I decided to change the protocol used to detect *lacZ* expression. Traditionally, X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), an organic compound made from galactose linked to indole, is cleaved by β-galactosidase (enzyme product of the *lacZ* gene) and the products of this reaction, in the presence of ferric and ferrous ions, form the blue precipitate that is scored in *lacZ*-reporter gene experiments. A more sensitive approach to detecting *lacZ* expression/activity is achieved by using Salmon-gal (S-gal; 6-chloro-3-indolyl-β-D-galactopyranoside) in conjunction with the tetrazolium salt NBT (nitroblue tetrazolium) (Sundararajan et al., 2012). The sensitivity of the S-gal approach to detecting *lacZ* activity can also easily lead to high background; therefore the amount of time that embryos were left in the S-gal + NBT solution was determined empirically (Fig. 3; see Materials and Methods).

Embryos assayed for *lacZ* expression from both lines indicated that the limb-specific expression pattern regulated by *M280* spans a narrow window of mouse embryonic development, from E10.5 to E11.5 (Figs. 4 and 5). In both lines, limb bud expression was first detected at E10.5 and only persisted through to E11.5. In all E12.5 *M280*-GWtg embryos collected I never detected *lacZ* expression in the developing limb. The shape of the expression pattern in E10.5 limb buds presaged the somewhat triangular shape of the E11.5 expression pattern. Moreover, *lacZ* expression was just as intense at E10.5 as it was at E11.5. The intensity of the signal in S-gal stained embryos also permits one to appreciate that the expression pattern appears to be mesenchymal and that it does not extend into ectodermal cells (such as the AER; this was later confirmed by sectioning through *M280*-GWtg limb buds; Figs. 3 and 6). The narrow window of *M280*-regulated, limb-specific expression implies that any aberrations in limb development, specifically in digits II-VI, in mice lacking *M280* would likely be the result of gene misregulation early in the mesenchymal cells that give rise to these autopod elements.

Figure 3. Salmon-gal staining technique preferred to detect *lacZ* expression in *M280*L1 and *M280*L2. (A-C) E11.5 *M280*L1 embryos. (D-F) E11.5 *M280*L2 embryos. (A, D) Reporter gene expression assayed with X-gal. (B, E) Reporter gene expression assayed with Salmon-gal; incubated overnight at room temperature. (C, F) Reporter gene expression assayed with Salmon-gal; incubated for 30 minutes at 37° C. Red rectangles indicate conditions selected for all subsequent analyses of *M280*L1 and *M280*L2. Scale bars = 0.5 mm.

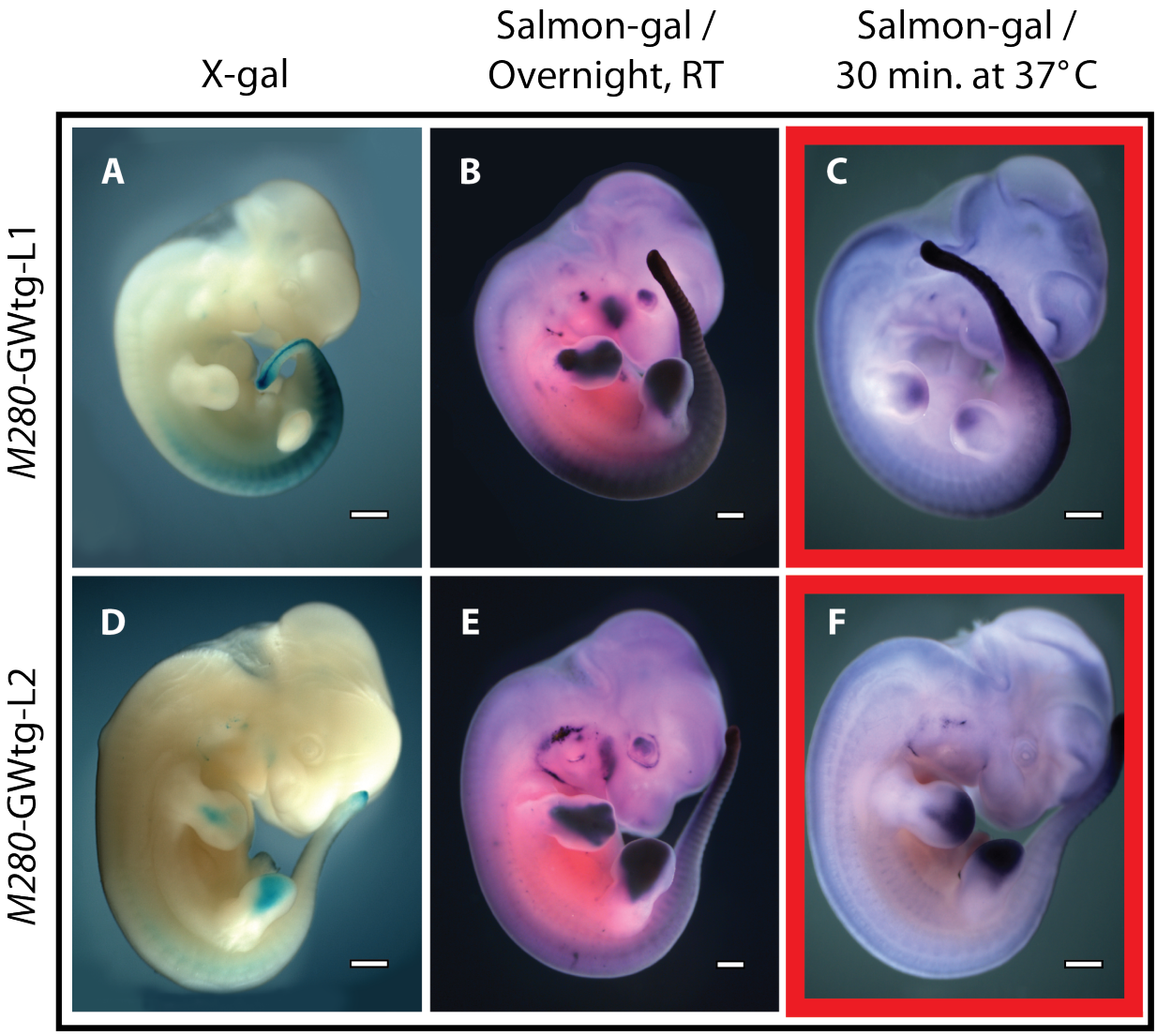

X-gal
Salmon-gal / Overnight, RT
Salmon-gal / 30 min. at 37°C
M280-GWtg-L1
M280-GWtg-L2
A
B
C
D
E
F

Figure 4. *M280*L1 drives *lacZ* expression in E10.5 and E11.5 limb buds. (A-D) Control embryos. (E-H) *M280*L1 embryos. (A, E) E9.5. (B, F) E10.5. (C, G) E11.5. (D, H) E12.5. Scale bars = 0.5 mm.

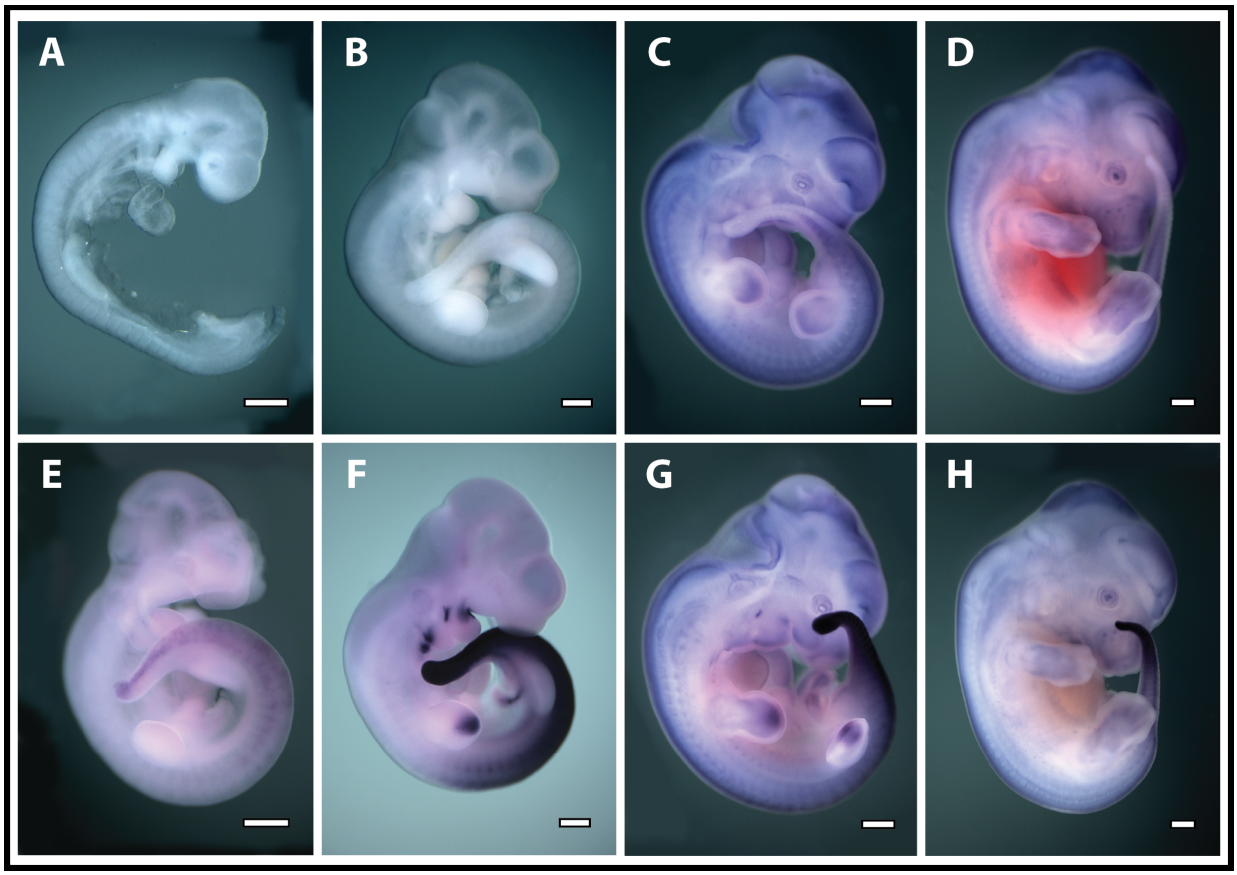

A
B
C
D
E
F
G
H

Figure 5. *M280*L2 drives *lacZ* expression in E10.5 and E11.5 limb buds. (A-E) Control embryos. (F-J) *M280*L2 embryos. (A, F) E9.5. (B, G) E10.5. (C, H) E11.5. (D, I) E12.5. (E, J) E13.5. Scale bars = 0.5 mm.

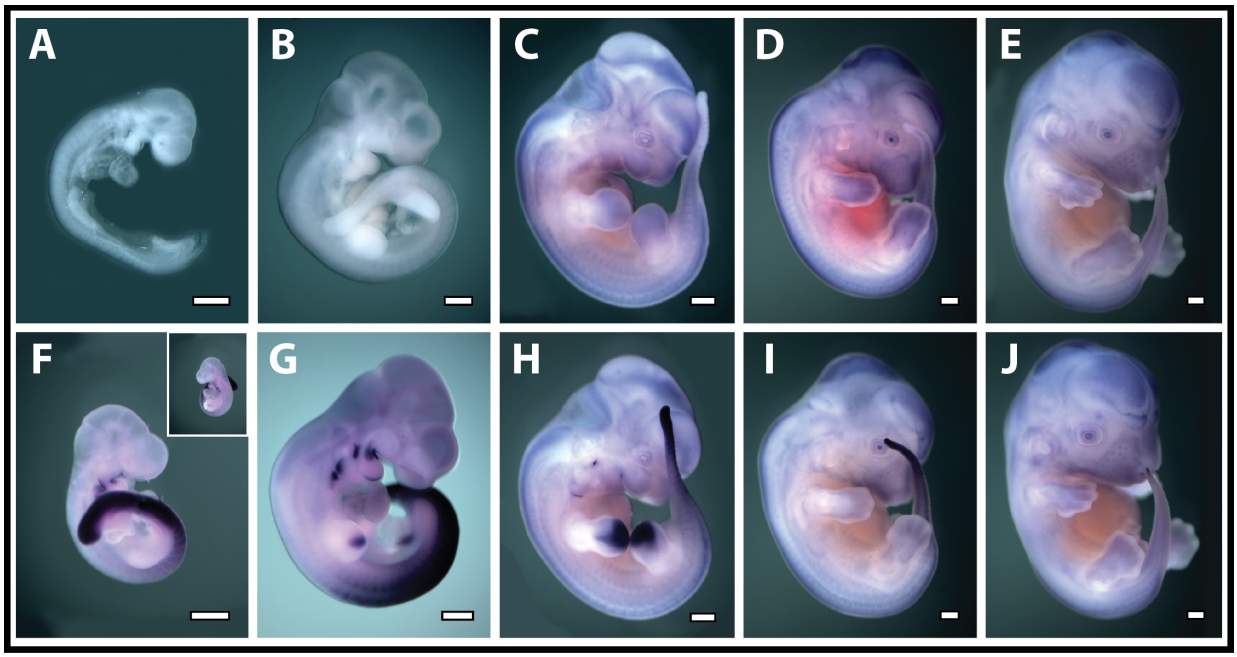
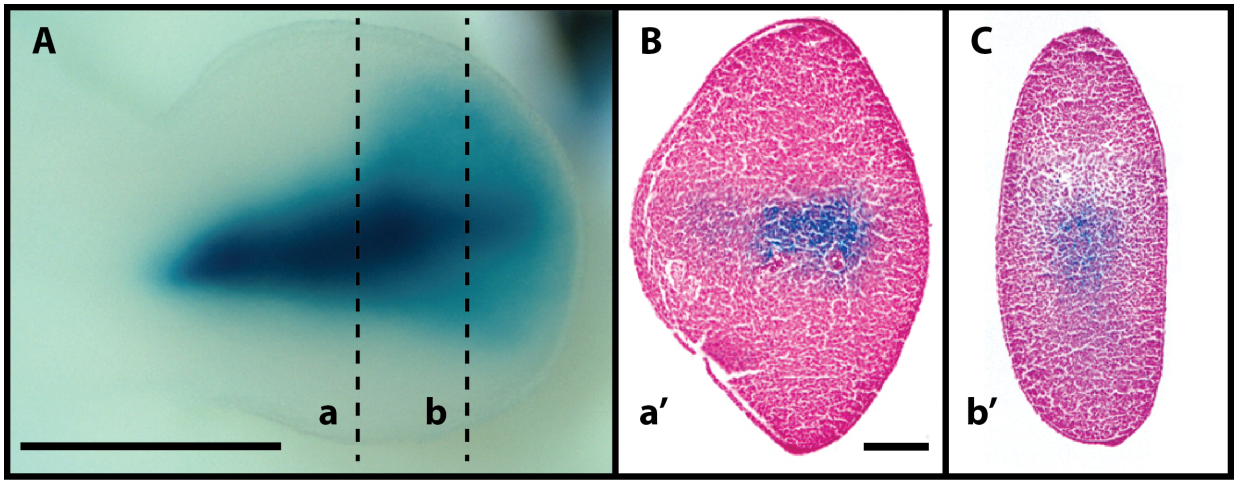

A
B
C
D
E
F
G
H
I
J

Figure 6. Cross-section through E11.5 *M280*-GWtg limb bud. (A) E11.5 forelimb bud. (B, C) Sections through E11.5 limb bud. Dotted lines (a) and (b) Show approximate section planes shown in (a') and (b'). *M280* is capable of driving *lacZ* expression in mesenchymal cells, but not ectodermal cells. *M280*-driven expression is strongest in the center of the forelimb bud, corresponding to the future location of digit III. Scale bar in (A) is 0.5 mm. Scale bar in B applies to C as well, and is 100 μm.


A
a
b
B
a'
C
b'

Expression of *lacZ* was also detected in the posterior neural tube, somites and tail bud as early as E9.5 (Figs. 4 and 5). Expression in the tail bud is maintained until E12.5, and occasionally at the very tip of the tail in E13.5 embryos. The most intense *lacZ* expression in the pharyngeal arches occurs at E10.5, and is clearly intense in arches 1-3, particularly in the oral groove of arch 1. The expression in the arches is weaker at E11.5, and concomitant with the development of more advanced facial structures at E12.5 the expression in the craniofacial region is absent or weak embryos of this age.

Histological sections of *M280* transgenic mice

To gain a better understanding of the spatial extent of the expression pattern regulated by *M280*, we sectioned numerous forelimb buds from E11.5 *M280*GWtg embryos. Sections were then stained with Nuclear Fast Red to distinguish *lacZ*-expressing cells from non-expressing cells. (Expression of *lacZ* in the sectioned embryos was originally assayed using X-gal rather than S-gal because at the time of sectioning we had not decided to use S-gal). Cells positive for *lacZ* were observed in the central interior of the E11.5 limb buds (Fig. 6). Expression was especially intense in the core group of expressing cells and weakened gradually, in gradient fashion, at increasingly greater distances from the core expressing cells. Cells positive for *lacZ* were not observed in the ectoderm, lending credence to the observation in over-exposed, S-gal stained *M280*GWtg embryos that *M280* does not regulate gene transcription the AER. I did not section transgenic embryos from other time points because *M280* is active in the limb during such a narrow developmental window that it did not seem necessary. The *lacZ* expression pattern noted in the sections corresponds to that detected in whole transgenic embryos and further supports the hypothesis that mice lacking *M280* could potentially manifest defects in the

development of the central autopod. Specifically, tissues corresponding to digits II-IV, but especially digit III, may be adversely affected in a developing limb bud that lacks *M280*.

Gross analysis of *M1442* transgenic mice

As was the case with *M280, M1442* was cloned upstream of the minimal promoter, *HSP68*, and the reporter gene, *lacZ*, using traditional cloning methods. The linearized and purified transgene was then injected into F_2 zygotes from (C57BL/6 X SJL)F_1 intercrosses and transferred to pseudopregnant Swiss females via oviduct transfer. As a first pass to determine whether *M1442* behaved similarly to *hs1442,* I collected E11.5 embryos (G_0 embryos) from foster mothers and scored them for the presence of the *M1442* transgene via *lacZ* staining (Fig. 2H).

Only a single embryo positive for this *M1442* transgene displayed *lacZ* expression in the E11.5 limb, and the pattern was somewhat different from that seen in *hs1442*-transgenic mice from the VEB (Fig. 2G, H). The *M1442* transgenic embryo manifested a *lacZ* expression pattern that ran along the anterior flank of the forelimb, with the highest intensity near the proximal margins of the limb bud and the weakest expression about midway along the anterior margins of the autopod. No expression was seen in the distal margins of the autopod, nor was expression detected in the posterior margins of the autopod as had been the case for ten of the eleven embryos expressing *lacZ* under control of *hs1442* as reported on the VEB. A similar, though fainter, expression pattern was seen in the hindlimbs of the *M1442* transgenic embryo. Portions of the frontonasal complex, forebrain, midbrain and numerous anterior somites also showed *lacZ* expression; a single embryo on the VEB also shows forebrain *lacZ* expression. Therefore, though my single *M1442* transgenic embryo did not compare extensively with *hs1442* transgenic embryos, it did demonstrate that *M1442* could at least direct gene expression in

E11.5 limb buds. Moreover, I was not willing to draw meaningful conclusions from a single embryo – the somewhat novel expression pattern in the anterior region of the limb bud could have been the result of the site of transgene insertion or it could have been the result of enhancer sequence differences between mouse and human. By the time this single transgenic result was available I had already made additional *M1442* transgene constructs using the Invitrogen Gateway Technology to more faithfully mimic the transgene vectors used on the VEB. Consequently, we generated G_0 transgenic embryos using constructs composed of *M1442* and the Gateway reporter construct described above (*M1442*-GWtg).

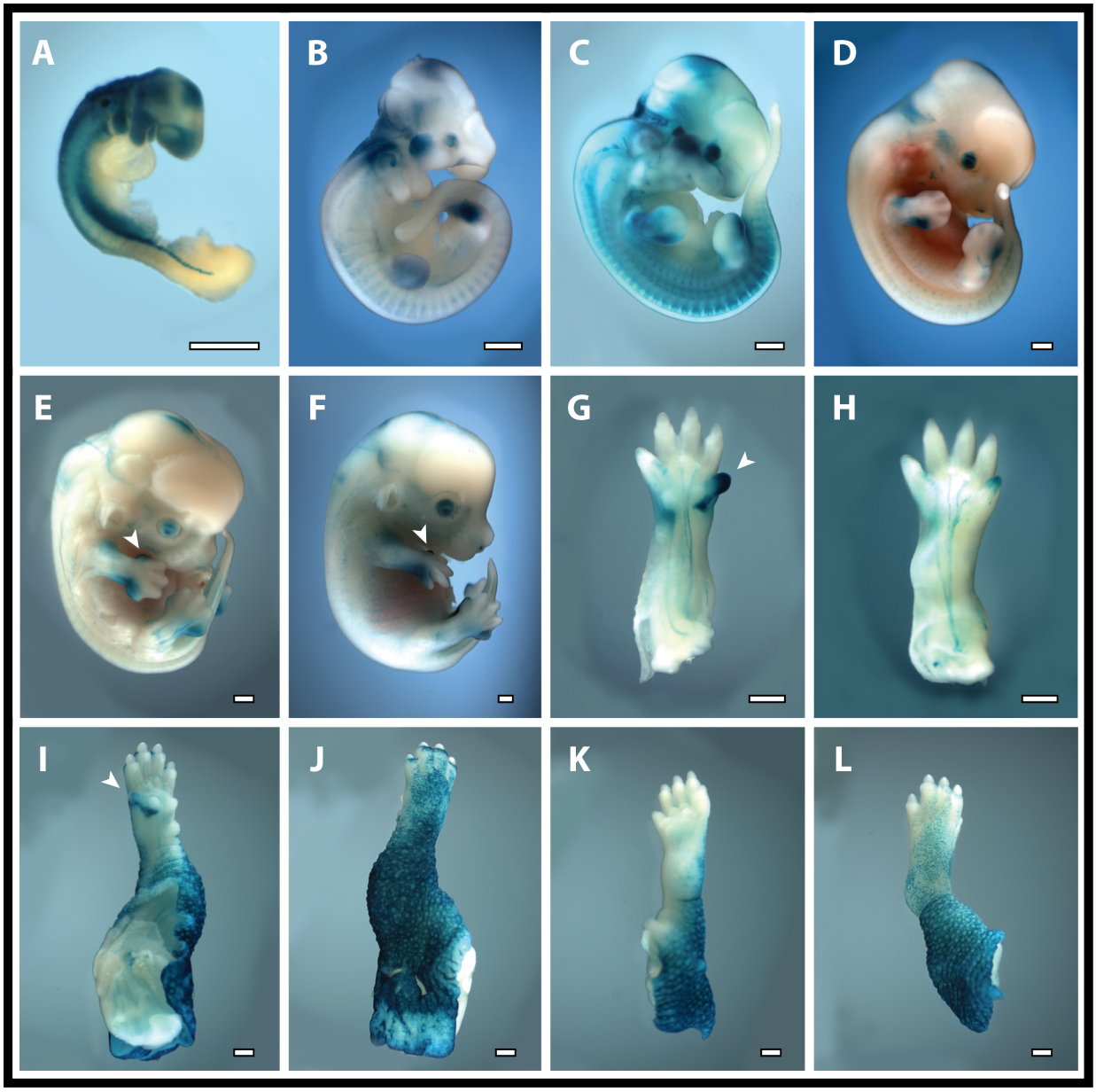
As was done for the original *M1442*-transgenic embryo, *M1442*-GWtg embryos were initially collected at E11.5 (G_0) and scored for their *lacZ* expression patterns in limb buds to determine whether the *M1442* regulated a well-defined, limb-specific expression pattern in the context of the Gateway transgene similar to that observed in *hs1442*-transgenic embryos. Positive *lacZ* expression was scored in four embryos from this experiment; all embryos manifested an identical limb bud expression pattern, though in one of the four the signal was very weak. Similar to *hs1442*-transgenic embryos, *M1442*-GWtg embryos had intense *lacZ* expression in the posterior margins of the autopod in both fore- and hindlimbs (Fig. 2I-L). However, this posterior expression was less graded towards the middle and anterior regions of the bud than that seen in numerous *hs1442*-transgenic embryos on the VEB. Curiously, as was seen in the single *M1442*-transgenic embryo described previously (non-Gateway) the forelimbs of G_0 *M1442*-GWtg embryos indicated active *lacZ* expression in the proximal-anterior borders of the limb bud (Fig. 2K, L). Again, this pattern was observed only in forelimbs, and was not observed in any *hs1442*-transgenic embryos on the VEB. It is very likely that the *M1442*-GWtg inserted into different locations in the four embryos. Therefore, the proximal, anterior *lacZ* expression patterns are not likely due to the random regulatory inputs at the sites of transgene integration, rather they may be due to sequence differences between the mouse and human

enhancers. *M1442*-GWtg embryos also expressed *lacZ* in the genital tubercle; this was noted in nine of eleven *hs1442* embryos on the VEB. One of the four *M1442*-GWtg embryos also showed expression in the developing eye. These observations permitted me to conclude that *M1442* is capable of driving gene expression in the posterior and anterior margins of E11.5 limb buds. With this information in hand I decided to generate stable *M1442*-GWtg lines to assess the regulatory capacity of *M1442*, specifically in the limb, throughout development.

I designed a primer set to specifically detect the presence of the *M1442*-GWtg. Of eight *M1442*-GWtg positive mice, two produced embryos that had detectable *lacZ* expression in E11.5 limbs. Thus, these two lines were used to characterize *M1442*'s regulatory capacity throughout development (*M1442*-GWtg-Line 1 and Line 2, or just *M1442*L1 and *M1442*L2). I will describe the expression pattern born by *M1442*-GWtg in each line in turn, beginning with *M1442*L1.

At E9 in *1442*L1 there is *lacZ* expression almost throughout the whole embryo, but it is particularly intense in the flanks of the embryo about a third of the distance from the developing tail bud, just ventral to the somites, in what is likely the lateral plate mesoderm that gives rise to the mesenchymal portion of the limb bud (Fig. 7A). By E10.5, reporter gene expression is weakly expressed in the limb bud in a diffuse manner, with a slightly more intense signal coming from the bud's posterior half (Fig. 7B). Similar to the patterns seen in previously described *M1442*-based transgenic embryos, by E11.5 the intense posterior signal of *lacZ* expression is detectable in fore- and hindlimb buds (Fig. 7C). Also visible at this time point is expression in the now expected proximal-anterior borders of the forelimb bud – though this signal is less intense than that of the posterior limb bud. By E12.5 it is clear that the posterior autopod expression is confined to the regions just proximal to and including the mesenchymal condensation that will give rise to digit V (Fig.7D). Additionally, the proximal, anterior expression has intensified. Also noteworthy is the fact that there is no intense proximal, anterior

Figure 7. *M1442*L1 drives *lacZ* expression in the limb throughout development. (A-F) E9.5-E14.5 embryos, respectively. (G, H) Dorsal view of E15.5 forelimb and hindlimb, respectively. (I-J) Dorsal and ventral view E16.5 forelimb, respectively. (K, L) Dorsal and ventral view of E16.5 hindlimb, respectively. Strong *lacZ* expression begins in the anterior forelimb at E13.5 and continues in the region of digit I until E16.5 (white arrowheads). Scale bars = 0.5 mm.


A
B
C
D
E
F
G
H
I
J
K
L

lacZ expression in the hindlimb at this point. In E13.5 forelimbs the posterior expression pattern is clearly confined to the base of future digit V and tissue just proximal to it – tissue that likely corresponds to posterior wrist elements (Fig. 7E). Importantly, at E13.5, the anterior *lacZ* expression is now detectable more distally than it was before, being present in developing digit I. In E13.5 hindlimbs one can detect expression in developing digit V, as well as a faint indication that anterior *lacZ* expression is present as a thin line running along the anterior border where the ventral and dorsal domains of the developing limb meet. Embryonic day 14.5 *lacZ* positive limbs are similar to E13.5 (Fig. 7F). By E15.5, expression is intense throughout forelimb digit I and in the wrist region just below the base of digit V (Fig. 7G, H). Hindlimbs at this time point are characterized by relatively intense *lacZ* expression in the tip of digit I. Expression is diminished greatly in digit I of forelimbs by E16.5, and is all but absent in the hindlimb footplate (Fig. 7I-L). Interestingly, expression is present in the skin throughout the entire limb and embryo in general at E16.5. Indeed, if the skin is removed one can see that there is very little expression in the underlying tissue (sections of *M1442*-GWtg animals also show that *M1442* can direct expression in the dermis of E15.5 embryos; Fig. 8). At E17.5 *lacZ* expression is absent in the limb (not shown). Thus, *M1442*L1 embryos demonstrate that *M1442*'s regulatory capacity is dynamic, active in specific, but changing locations of the developing limb, with a strong activity in the anterior and posterior margins of the limb bud.

In E11.5 embryos from the *M1442*L2 line, *lacZ* expression is intense in both the anterior and posterior margins of the forelimb. In the hindlimb, though the expression is strongest in the posterior bud, there is a patch of relatively strong expression near its anterior base (Fig. 9A). Reporter gene expression persists in this manner in the limb bud through to E14.5 (Fig. 9B-D). The intensity of the *lacZ* signal decreases in the anterior border of the forelimb as development progresses on until it is concentrated largely into digit I; a similar process occurs in the hindlimb, with intense anterior expression confined to the distal tip of digit I by E15.5 (Fig. 9E,

Figure 8. E16.5 *M1442*-GWtg expression confined largely to skin. (A) E16.5 control embryo. (B) E16.5 *M1442*-GWtg embryo. (C, E) Dorsal view of skinned *M1442*L1 and *M1442*L2 forelimbs, respectively. (D, F) Ventral view of skinned *M1442*L1 and *M1442*L2 forelimbs, respectively. Note the strong expression that still remains in the region of the falciform carpal, proximal to digit I. (G) Section through digits of an E15.5 *M1442*-GWtg forelimb showing *lacZ* expression confined to region corresponding to developing dermis. All scale bars = 0.5 mm, except in (G) wherein scale bar = 25μm.

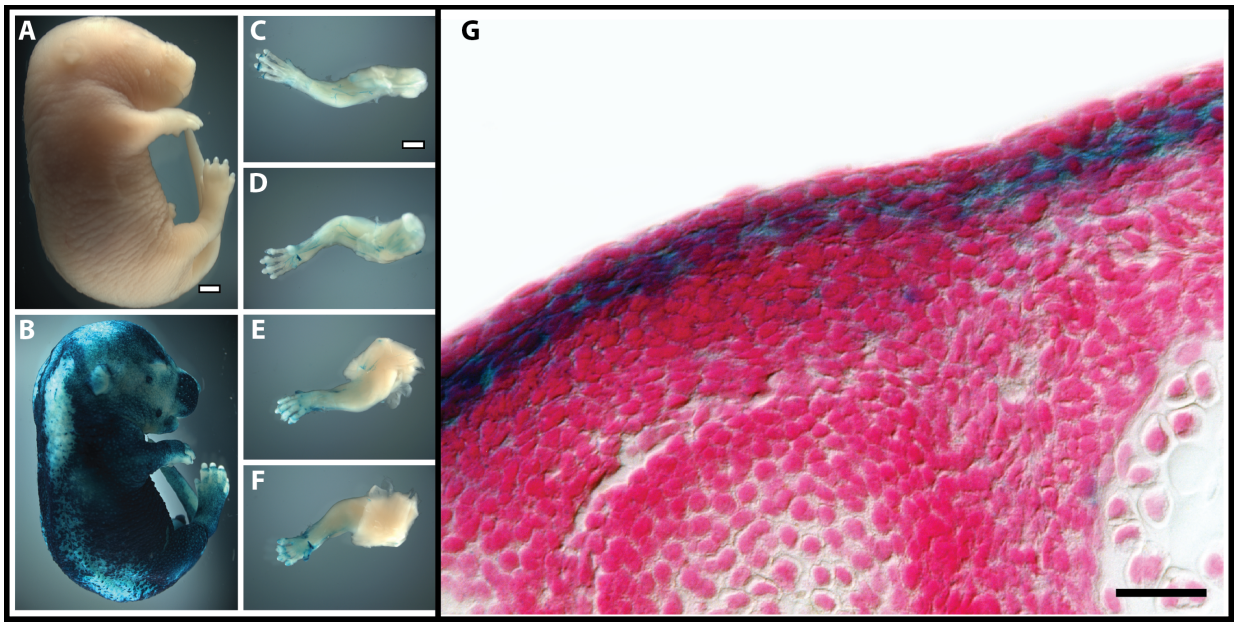
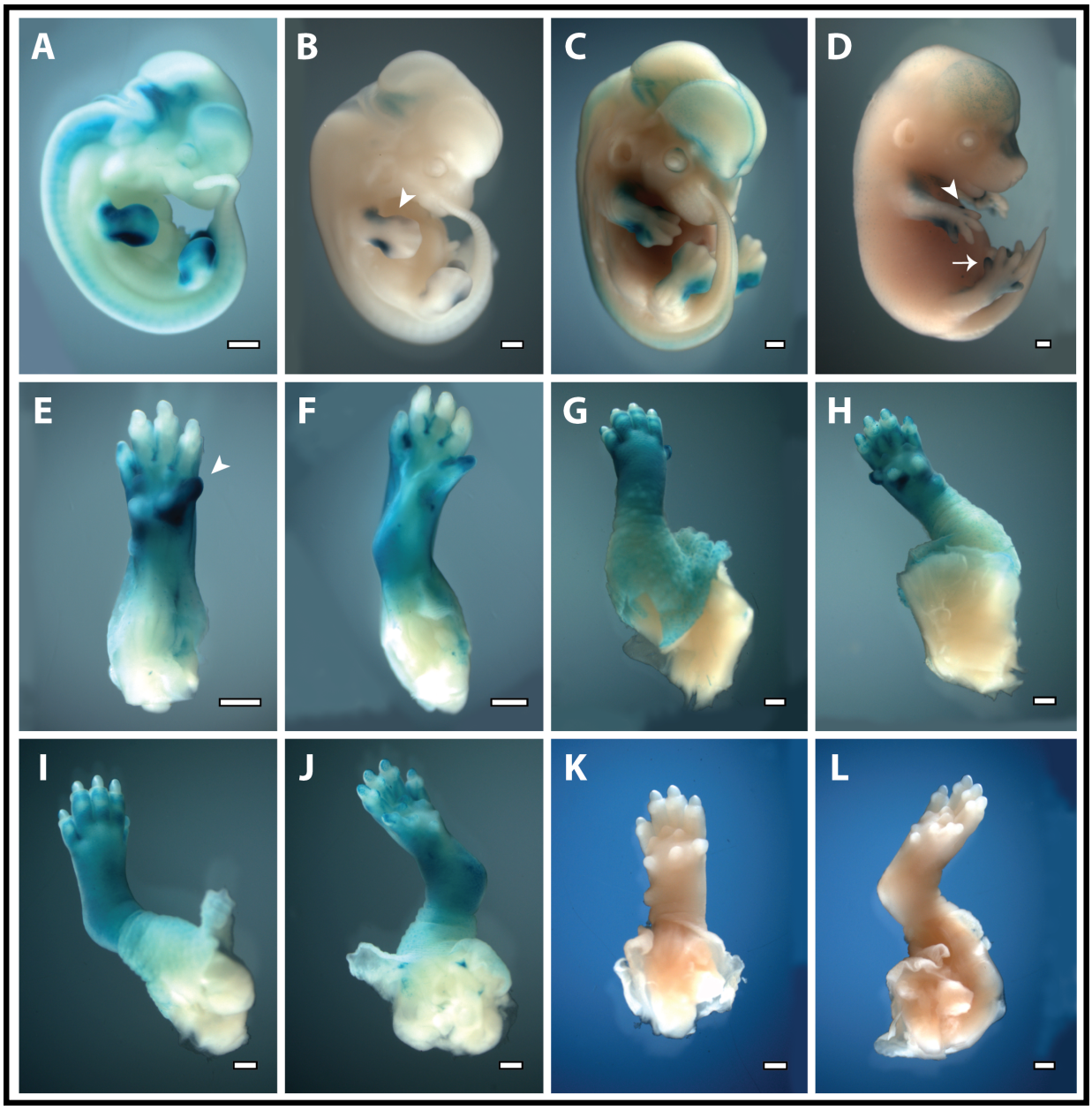

A
B
C
D
E
F
G

Figure 9. *M1442*L2 drives *lacZ* expression in the limb throughout development. (A-D) E11.5-E14.5 embryos, respectively. (E, F) Dorsal view of E15.5 forelimb and hindlimb, respectively. (G, H) Dorsal and ventral view E16.5 forelimb, respectively. (I, J) Dorsal and ventral view of E16.5 hindlimb, respectively. (K, L) Ventral view of E17.5 *M1442*L2 positive forelimb and hindlimb, respectively. Strong *lacZ* expression begins in the anterior forelimb at E14.5 and continues in the region of digit I until E16.5 (white arrowheads). Anterior expression in the hindlimb begins at E14.5 but does not persist in digit V beyond E15.5. *M1442*L2-mediated limb expression ceases by E17.5. Scale bars = 0.5 mm.


A
B
C
D
E
F
G
H
I
J
K
L

F). Also, in E15.5 forelimbs there appears to be expression in the developing epidermis, and perhaps the underlying tissues as well, near the base of digits II-IV. By E16.5, *M1442*L2 forelimbs have lost their intense *lacZ* expression in digit 1 and expression has largely moved to the skin of the entire limb (and the whole embryo in general). As was the case for *M1442*L1, when the skin is removed from E16.5 limbs there is very little *lacZ* expression in the underlying tissue (Fig. 9G-H; Fig. 8). Additionally, at E16.5, digit I in the hindlimb displays no intense, unique *lacZ* signal relative to the other digits (Fig. 9I, J). By E17.5, *M1442* ceases to regulate gene expression in the fore- and hindlimbs and in the skin of the entire embryo (Fig. 9K, L).

Comparing the expression patterns of both *M1442*L1 and *M1442*L2 embryos permits a consensus picture of *M1442*-regulated, limb-specific gene expression to crystalize. From the observations above, I concluded that *M1442* is capable of directing expression in the limb bud as soon as it begins developing – its activity, though diffuse, obviously exerted at E10.5, but also perhaps as early as E9 in the lateral plate mesoderm that gives rise to the limb. Furthermore, *M1442* is capable of directing gene expression to the anterior and posterior regions around digits I and V as early as E11.5 (in Line 2; though digit I expression begins at E13.5 in Line 1) and persists until E15.5. By E16.5, digit-specific and wrist/ankle-specific expression regulated by *M1442* has diminished and regulatory capacity has switched to direct expression in the skin. Guided limb-specific gene expression by *M1442* stops at E17.5. Therefore, phenotypes associated with deleted *M1442* would likely result in aberrations in digits I and V, and in the wrist/ankle elements at the base of these digits, in particular the hamate, triquetral and pisiform bones proximal to digit V. Additionally, the *1442*GWtg expression patterns indicate that aberrations in mice lacking *M1442* could be detected in the developing skin. It is important to note too that on the posterior side of the limb, the *lacZ* expression pattern persists proximally all the way to the distal tip of the ulna; though *lacZ*-expressing cells do not appear in the ulnar tip itself there is clear expression in the flanking dermal and mesenchymal cells. Thus, it cannot be

ruled out that any necessary cellular crosstalk between the differentiating cells of the ulna and the surrounding tissue may also be disrupted in mice lacking *M1442*, in which case the morphology of either communicating tissue may be altered.

Histological sections of *M1442* transgenic mice

To gain a better understanding of the spatial extent of the expression pattern regulated by *M1442* I characterized *M1442*-driven gene expression in sectioned limbs from two time points: E11.5 and E15.5.

At E11.5 the *M1442*-GWtgs drive strong expression in the posterior margins of the distal forelimb bud (Fig. 10 A-C). The posterior *lacZ* signal is stronger than the anterior signal and extends deeper into the central mesenchymal cells. It is noteworthy that expression is not detected in the epidermal cells. From these observations it seems likely that *M1442* may contribute to the development of distal-posterior structure, such as digits IV and V, as well digit I in the anterior limb bud.

Serial sections through E15.5 *M1442*GWtg forelimbs revealed a somewhat complex *lacZ* expression pattern, one that interfaced with developing bone elements, likely dermal cells, and soft tissue surrounding both digits and wrist bones (Fig. 11A-E). Tracking the position of E15.5 digits through a series of sections is easy enough, but the task of identifying carpal (wrist) elements is a bit trickier. Numerous carpal bones have irregular shapes in cross-section and can roughly span the entire dorsal-ventral axis of the limb, seemingly changing position from one section to the next. To aid in proper identification of carpal bones I compared my cross-sections to the color-coded, 3D-reconstructions of mouse limbs found at the EMAP eMouse Atlas Project (http://www.emouseatlas.org) (Delaurier et al., 2008; Husz et al., 2012).

Figure 10. Cross-section through E11.5 *M1442*-GWtg limb bud. (A) Left E11.5 forelimb bud. (B, C) Sections through E11.5 limb bud. Dotted lines (a) and (b) Show approximate section planes shown in (a') and (b'). *M1442* is capable of driving *lacZ* expression in mesenchymal cells, but not ectodermal cells. *M1442*-driven expression is strongest in the posterior limb, corresponding to the future location of digit V. Scale bar in (A) is 0.5 mm. Scale bar in B applies to C as well, and is 50 μm.

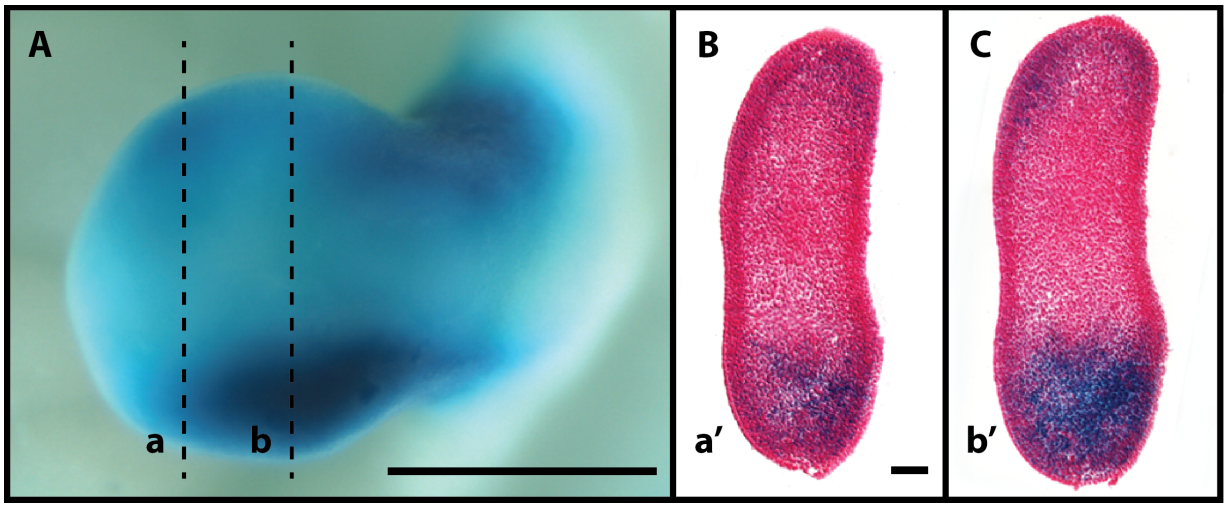
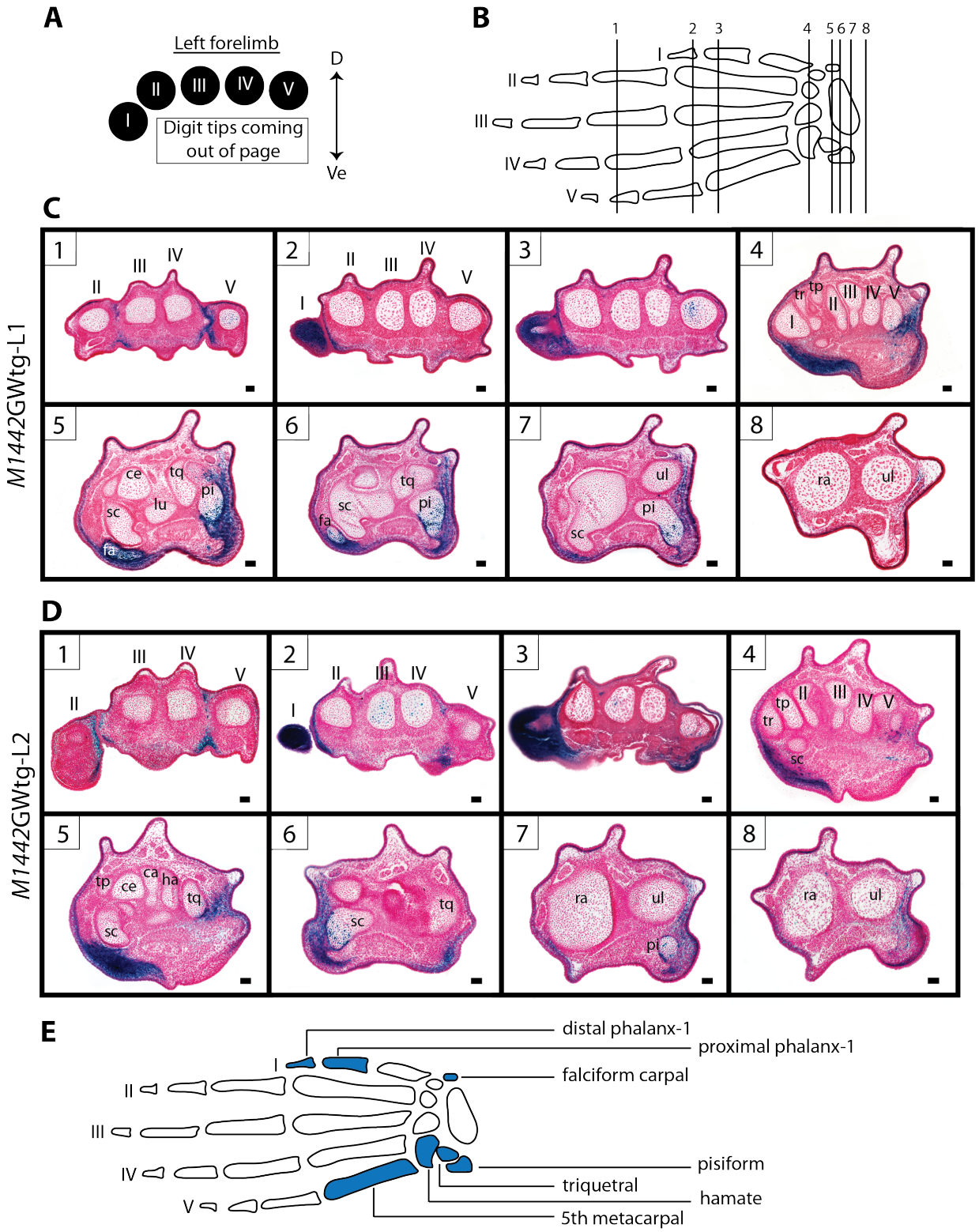

A
a
b
B
a′
C
b′

Figure 11. Cross-sections through E15.5 *M1442*-GWtg limb buds. (A) Orientation guide for cross-sections. (B) Schematic of left autopod; Arabic numerals and associated vertical lines correspond to approximate plane of sections shown in (C) and (D). Boxed Arabic numerals in (C) and (D) correspond to vertical lines in (B). (C) Cross-sections through E15.5 *M1442*L1 forelimb. (D) Cross-sections through E15.5 *M1442*L2 forelimb. (E) Summary diagram of autopod elements that manifest *lacZ* signal in both *M1442*-GWtg lines. Roman numerals indicate digit rays. Abbreviations in (A): D = dorsal, Ve = ventral. Other abbreviations: ce = centrate, fa = falciform carpal, ha = hamate, lu = lunate, pi = pisiform, ra = radius, sc = scaphoid, tr = trapezium, tp = trapezoid, tq = triquetral, ul = ulna. Scale bars = 50 μm.


A
Left forelimb
I
II
III
IV
V
Digit tips coming out of page
D
Ve
B
1
2
3
4
5
6
7
8
I
II
III
IV
V
C
M1442GWtg-L1
ce
tq
pi
sc
lu
fa
tp
tr
ra
ul
D
M1442GWtg-L2
ca
ha
E
distal phalanx-1
proximal phalanx-1
falciform carpal
pisiform
triquetral
hamate
5th metacarpal

In sections of E15.5 forelimbs, *lacZ* signal is occasionally observed in the distal tip of digit V, in the third phalanx. This expression does not persist along the length of digit V, but does reappear in the fifth metacarpal (Fig. 11C.3, D.3). It appears to be absent beyond the distal-most phalanx of digit V. Along the lengths of digits II and III, and digits IV and V, *lacZ* expression does appear to be relatively constant in what looks like the developing dermal layer of the skin – this perhaps foreshadows the almost exclusive epidermal expression pattern driven by *M1442* observed in E16.5 limbs (Fig. 11C.1-3, D.1-3). All distal tissues comprising digit I are positive for *lacZ* expression, but curiously the first metacarpal and carpal elements just proximal to the first metacarpal – the trapezium in particular – do not express *lacZ* (Fig. 11C.2, 3 and D.2, 3). The more ventral carpal bone known as the falciform carpal is, however, *lacZ* positive (Fig. 11C.5, D.5). This indicates that on the anterior side of the limb, strong *M1442* expression is restricted to digit I and the falciform carpal. In some sections containing cross-sections of the fifth metacarpal one can see positive *lacZ* signal in the bone's interior. Also, in one of my transgenic lines, *M1442*GWtgL2, *lacZ* expression can be detected in the base of the third and fourth metacarpals (Fig. 11D.2).

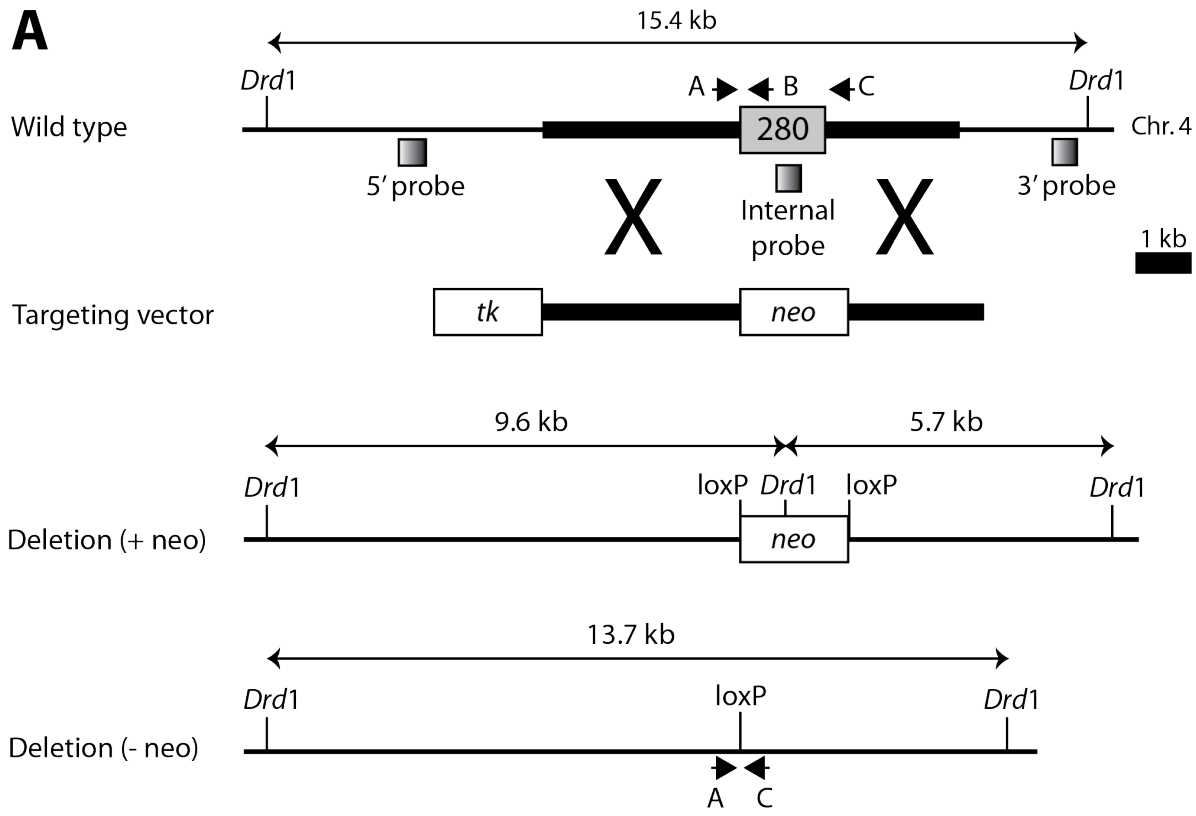
Proximal to the digits, several carpal elements manifest *lacZ* expression on the posterior side of the limb. The lateral, or posterior-most half of the triquetral bone, which is positioned below digit V, shows strong X-gal staining, as does the more ventrally located pisiform bone (Fig. 11C.5-7 and D.5, 7). Interestingly, while the *lacZ* expression in the triquetral appears most commonly in the interior trabecular bone cells, the expression in the pisiform is found both in the interior trabecular cells and in the thin layer of compact cells surrounding the spongy interior (Fig. 11C.6, D.7). Also, I cannot rule out that the posterior-most (or lateral-most) portion of the hamate bone does not also manifest *lacZ* expression because in particular sections distinction between the hamate and fifth metacarpal is not easily made (Fig. 11C.4, D.4). Therefore, it seems plausible to make a general assertion: *1442*-driven gene expression in the posterior

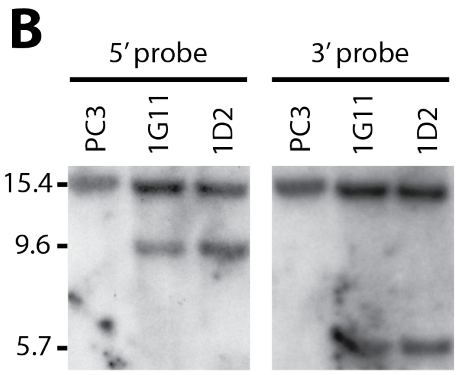
margins of the distal forelimb are restricted to the fifth metacarpal and to the lateral margins of all posterior carpal bones including the hamate, triquetral and the pisiform. The *lacZ* expression pattern noted in the sections corresponds to that detected in whole transgenic embryos and further supports the hypothesis that mice lacking *M1442* could potentially manifest defects in the development of digit I in the anterior autopod, and digit V in the posterior autopod. Additionally, carpal bones proximal to digit V, including the hamate, triquetral and pisiform bones would also be expected to manifest developmental abnormalities in mice lacking *M1442*.

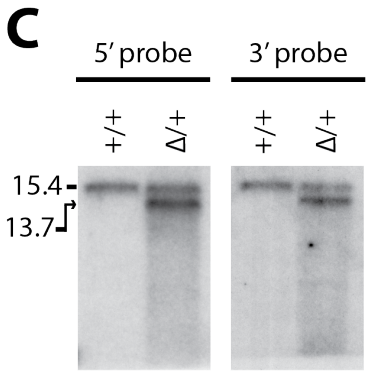
Generating *M280* Gene targeted mice

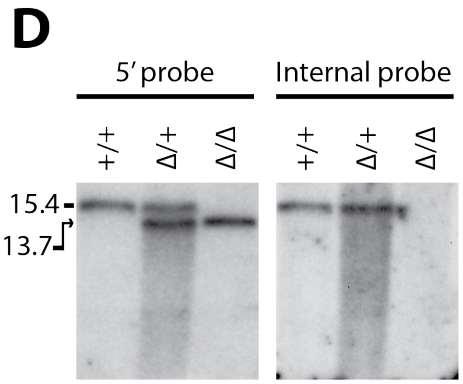
To functionally test the role of *M280* during mouse limb development we generated mice that lacked *M280* using traditional gene targeting methods (referred to as *M280*-knockouts or $M280^{tm1}$ mice; Material and Methods). Deletion of *M280* was carried out in mouse embryonic stem (ES) cells derived from a 129S6/SvEv genetic background. I designed Southern Blot probes that permitted me to verify the targeting event on both the 5'- and 3'-side of the *M280* locus (Fig. 12B). All chimeric *M280*-knockout mice generated from these cells were bred to 129S6/SvEv mice; therefore, all analyses described herein were conducted on 129S6/SvEv mouse tissue to avoid variation that may arise due to differences in genetic background. The *M280* construct was designed such that a *loxP*-flanked *neomycin resistance* gene (neo^r), used to aid in the selection of ES cells that underwent a targeting event, replaced the M280 enhancer. However, all progeny derived from male chimeric mice generated from the targeted ES cells would not possess the neo^r gene because the targeted ES cells also possessed the *Protamine-Cre recombinase* transgene (PC3 ES cells). This transgene is specifically activated during spermatogenesis, permitting the Cre recombinase protein to recombine any *loxP* sites present in the genome of sperm derived from male chimeras (O'Gorman et al., 1997). Therefore, all

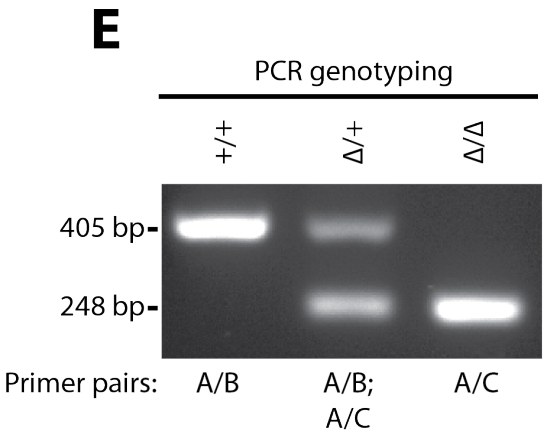
Figure 12. *M280* gene targeting strategy. (A) *M280* targeting strategy. Primers represented by black arrows; Southern Blot probes represented by shaded boxes. (B) Southern Blot showing targeting event in PC3 embryonic stem cells used to generated *M280* knockout mice. (C) Southern Blot showing removal of neo^r via Cre recombination of *loxP* sites. (D) Southern Blot for sequence internal to *M280* demonstrates its absence in $M280^{tm1}$ mice. (E) *M280* heterozygous and homozygous mice are distinguishable by PCR.


A
15.4 kb
Drd1
A
B
C
Drd1
Wild type
280
Chr. 4
5′ probe
Internal probe
3′ probe
1 kb
Targeting vector
tk
neo
9.6 kb
5.7 kb
Drd1
loxP
Drd1
loxP
Drd1
Deletion (+ neo)
neo
13.7 kb
Drd1
loxP
Drd1
Deletion (- neo)
A
C


B
5′ probe
3′ probe
PC3
1G11
1D2
PC3
1G11
1D2
15.4
9.6
5.7


C
5′ probe
3′ probe
+/+
Δ/+
+/+
Δ/+
15.4
13.7


D
5′ probe
Internal probe
+/+
Δ/+
Δ/Δ
+/+
Δ/+
Δ/Δ
15.4
13.7


E
PCR genotyping
+/+
Δ/+
Δ/Δ
405 bp
248 bp
Primer pairs:
A/B
A/B;
A/C
A/C

heterozygous, *M280*^{tm1} positive offspring obtained from male chimeras possessed a single *loxP* site (34 bps) in place of the *M280* enhancer (Fig. 12C).

Heterozygous *M280*^{tm1} offspring from 129S6/SvEv female X *M280*^{tm1} male chimera crosses were intercrossed to generate *M280*^{tm1} knockout (-/-) mice. The same Southern Blot probes used to verify the targeting event in the PC3 ES cells were used to validate the removal of the *neo*^{r} gene in DNA obtained from *M280*^{tm1} heterozygous animals. Additionally, I designed a Southern Blot probe complimentary to *M280* and showed that indeed this enhancer is absent in *M280*^{tm1} null mice (Fig. 12D). For continued ease in genotyping *M280*^{tm1} mice I also designed a set of PCR primers that can distinguish between wild-type, heterozygous and homozygous *M280*^{tm1} mice (Fig. 12E). In the wild-type situation one primer anneals to the sequence just 5' of *M280* while its cognate primer binds within *M280* itself; together they generate a 405 bp amplicon. An additional primer was designed to anneal to sequence 3' of *M280*; however, the distance between the 5' and 3' primers is too great for faithful amplicon production so that even in the wild-type condition only the 405 bp amplicon is generated. Alternatively, in the heterozygous or homozygous *M280*^{tm1} condition, *M280* is absent. Therefore the primer designed to bind within *M280* cannot do so and because the genome is shortened by the distance equal to the size of *M280* the 5'- and 3'-primers can generate a 248 bp amplicon.

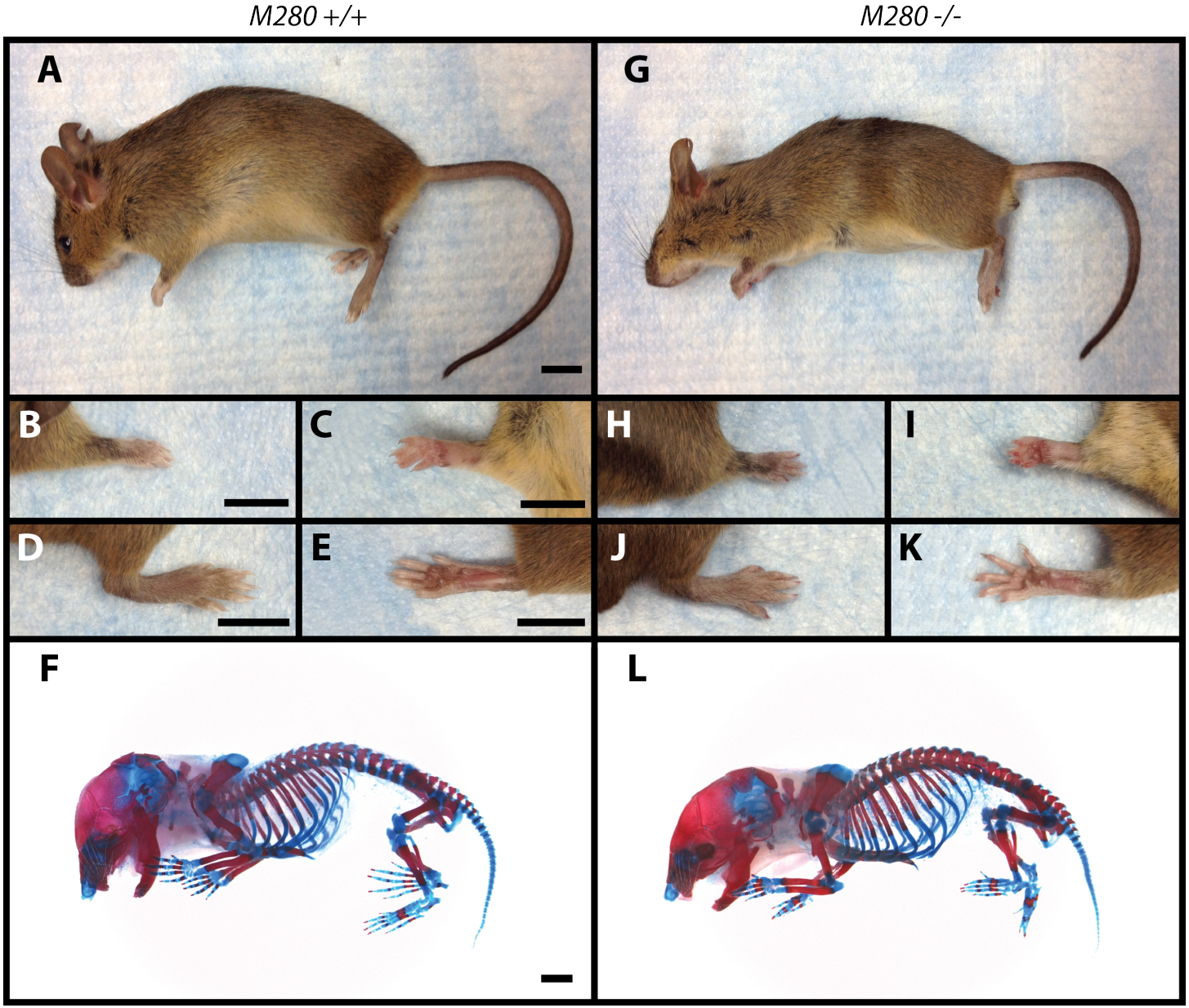
Heterozygous and homozygous *M280*^{tm1} mice are born in expected Mendelian ratios, are viable, fertile, and exhibit no obvious external limb phenotypes (Table 6; Fig. 13). To assess potential limb skeletal phenotypes that were not immediately available upon gross evaluation of the external limbs I stained newborn *M280*^{tm1} -/- pups with alizarin red and alcian blue. Alizarin red is a dye that stains bones red, and alcian blue is a dye that stains cartilage blue. Alizarin red and alcian blue stained newborn *M280*^{tm1} -/- pups were indistinguishable from wild-type littermates. In particular, relative bone position and morphology of fore- and hindlimbs dissected away from alizarin redand alcian blue stained *M280*^{tm1} -/- newborn pups appeared

Table 6. Chi-square test for expected Mendelian ratios in offspring from heterozygous *M280^{tm1}* intercrosses.

(N = 87)	WT	Het	Null
Expected:	21.75	43.5	21.75
Observed:	21	46	20

Two degrees of freedom
Chi-square = 0.31
Not significant at P = 0.5

Figure 13. Newborn and adult *M280* null mice are indistinguishable from control littermates. (A-F) Control animals. (G-L) *M280* null animals. (B, C, H, J) Dorsal and ventral forelimbs, respectively. (D, E, J, K) Dorsal and ventral hindlimbs, respectively. (F, L) Alizarin red and alcian blue stain for bone and cartilage in control (F) and mutant (L) newborn pups. Scale bars in (A-E, G-K) = 9 mm. Scale bars in (F, L) = 1 mm.


M280 +/+
M280 -/-
A
G
B
C
H
I
D
E
J
K
F
L

similar to wild-type limbs. Although observations from my *M280*GWtg embryos suggested that phenotypes would most likely arise in digits I, II or III, I noted no aberrations in the cartilage and bone elements of these digits (Fig. 13). These results were corroborated with histological sections of control and mutant E18.5 limbs (Figs. 14 and 15). Therefore, initial external and internal observations of limb morphology suggested that *M280* is not required for proper limb development.

Generating *M1442* Gene targeted mice

To functionally test the role of *M1442* during mouse limb development I generated mice that lacked *M1442* using traditional gene targeting methods (referred to as *M1442*-knockouts or $M442^{tm1}$ mice; Material and Methods; Fig. 16). Like *M280*, deletion of *M1442* was carried out in mouse ES cells derived from a 129S6/SvEv genetic background. I designed Southern Blot probes that permitted me to verify the targeting event on both the 5'- and 3'-side of the *M1442* locus within the ES cells (Fig. 16B). All chimeric *M1442*-knockout mice generated from these ES cells were bred to 129S6/SvEv mice; therefore, all analyses described herein were conducted on 129S6/SvEv mouse tissue to avoid variation that may arise due to differences in genetic background. The targeting methodology used at the *M1442* locus was the same as that used at the *M280* locus: the neo^r gene that replaced *M1442* in the targeting event was subsequently removed in the germline of *M1442*-deleted chimeras because I used the PC3 ES cell line for the *M1442* targeting event (O'Gorman et al., 1997). Therefore, all heterozygous, $M1442^{tm1}$ positive offspring obtained from male chimeras possessed a single *loxP* site (34 bps) in place of the *M1442* enhancer.

Heterozygous $M1442^{tm1}$ offspring from 129S6/SvEv female X $M1442^{tm1}$ male chimera crosses were intercrossed to generate $M1442^{tm1}$ knockout mice. I designed a PCR genotyping

Figure 14. Forelimb autopod elements present in E18.5 and newborn *M280* null embryos. (A-C) Control tissue. (D-F) *M280* mutant tissue. (A, D) Alizarin red and alcian blue stain of newborn forelimbs. (B, E) H&E stained planar sections through forelimb autopods. (C, F) H&E stained frontal sections through forelimb autopods. Abbreviations: ca = capitate; dp = distal phalanx; ha = hamate; ip = intermediate phalanx; pp = proximal phalanx; ra = radius; tp = trapezoid; ul = ulna. Roman numerals correspond to digit rays.

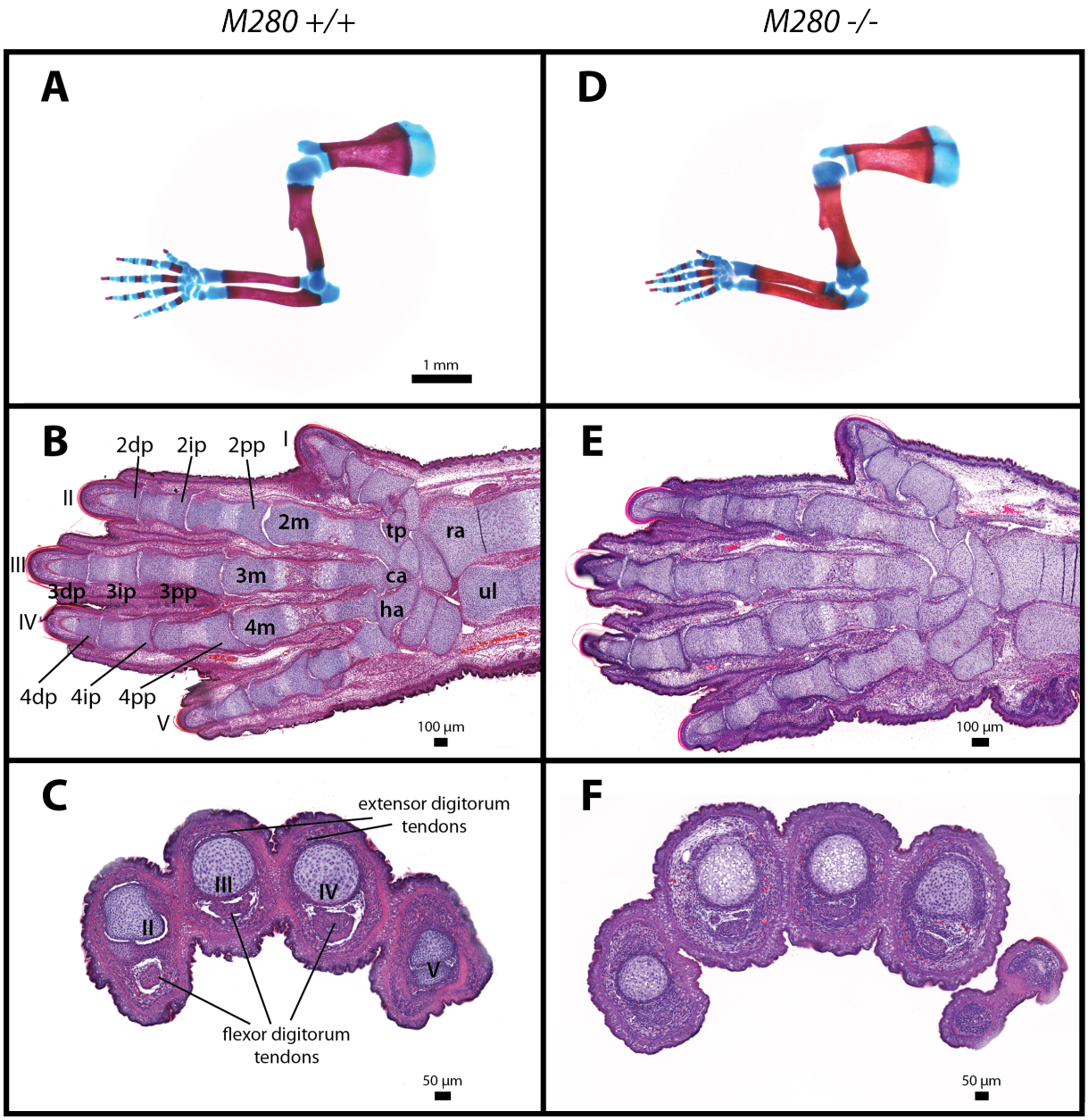

M280 +/+
M280 -/-
A
1 mm
D
B
2dp
2ip
2pp
I
II
III
IV
2m
tp
ra
3m
ca
ul
3dp
3ip
3pp
ha
4m
4dp
4ip
4pp
V
100 µm
E
100 µm
C
extensor digitorum
tendons
III
IV
II
V
flexor digitorum
tendons
50 µm
F
50 µm

Figure 15. Hindlimb autopod elements present in E18.5 and newborn *M280* null embryos. (A-C) Control tissue. (D-F) *M280* mutant tissue. (A, D) Alizarin red and alcian blue stain of newborn forelimbs. (B, E) H&E stained planar sections through hindlimb autopods. (C, F) H&E stained frontal sections through hindlimb autopods. Abbreviations: ca = calcaneus; cu = cuboid; dp = distal phalanx; ic = intermediate cuneiform; ip = intermediate phalanx; m = metatarsal; na = navicular; pp = proximal phalanx; ta = talus. Roman numerals correspond to digit rays.

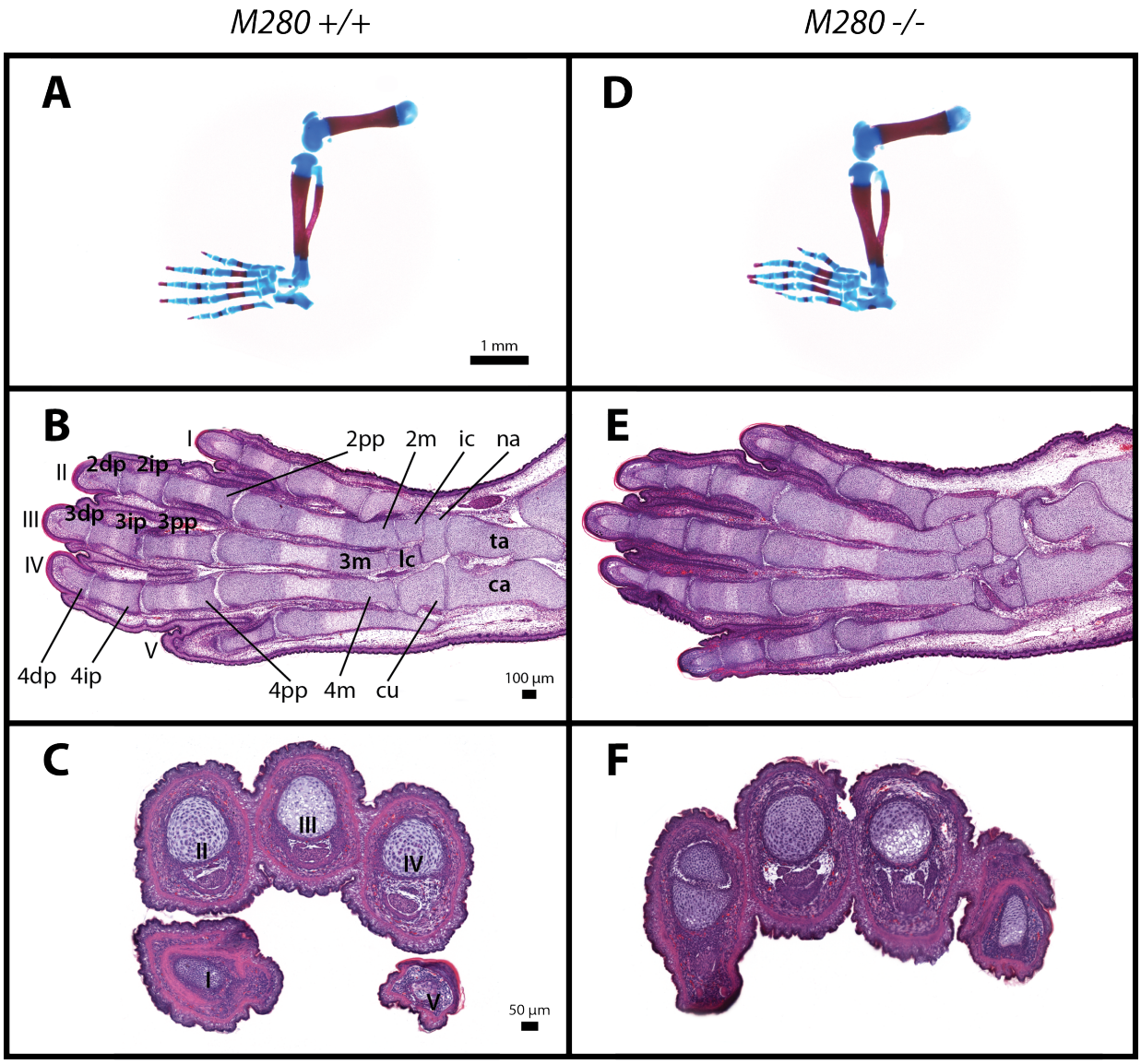
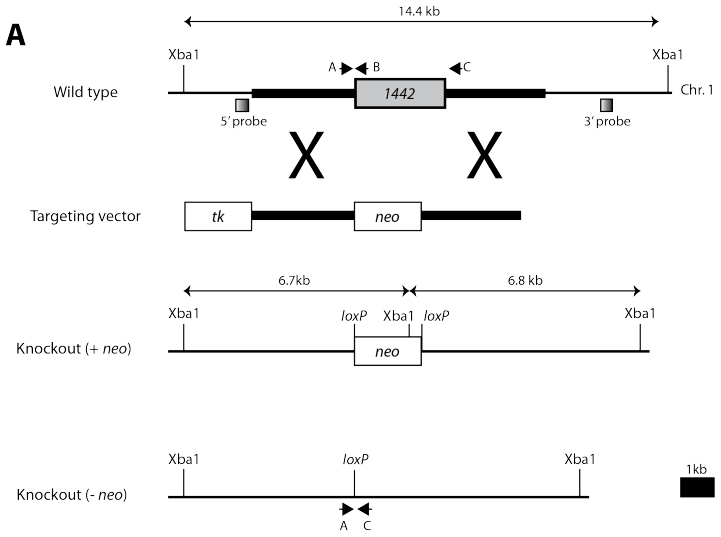
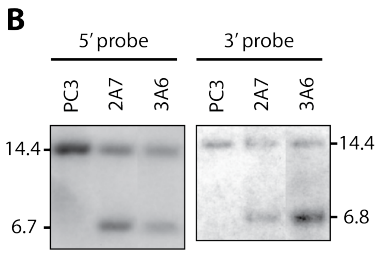
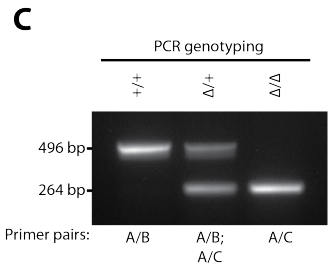

M280 +/+
M280 -/-
A
1 mm
D
B
I
II
III
IV
V
2dp
2ip
3dp
3ip
3pp
4dp
4ip
2pp
2m
ic
na
ta
3m
lc
ca
4pp
4m
cu
100 µm
E
C
II
III
IV
I
V
50 µm
F

Figure 16. *M1442* gene targeting strategy. (A) *M1442* targeting strategy. Primers represented by black arrows; Southern Blot proves represented by shaded boxes. (B) Southern Blot showing targeting event in PC3 embryonic stem cells used to generated *M1442* knockout mice. (C) *M1442* heterozygous and homozygous mice are distinguishable by PCR.


A
14.4 kb
Xba1
Xba1
A
B
C
Wild type
1442
Chr. 1
5′ probe
3′ probe
X
X
Targeting vector
tk
neo
6.7kb
6.8 kb
Xba1
loxP
Xba1
loxP
Xba1
Knockout (+ neo)
neo
Xba1
loxP
Xba1
1kb
Knockout (- neo)
A
C


B
5′ probe
3′ probe
PC3
2A7
3A6
PC3
2A7
3A6
14.4
6.7
14.4
6.8


C
PCR genotyping
+/+
Δ/+
Δ/Δ
496 bp
264 bp
Primer pairs:
A/B
A/B;
A/C
A/C

strategy to distinguish between heterozygous and homozygous *M1442*^{tm1} and wild-type mice (Fig. 16C). In the wild-type situation one primer anneals to the sequence just 5' of *M1442* while its cognate primer binds within *M1442* itself; together they generate a 496 bp amplicon. An additional primer was designed to anneal to sequence 3' of *M1442*; however, the distance between the 5' and 3' primers is too great for faithful amplicon production so that even in the wild-type condition only the 496 bp amplicon is generated. Alternatively, in the heterozygous or homozygous *M1442*^{tm1} condition, *M1442* is absent. Therefore the primer designed to bind within *M1442* cannot do so and because the genome is shortened by the distance equal to the size of *M1442* the 5'- and 3'-primers can generate a 264 bp amplicon.

Heterozygous and homozygous *M1442*^{tm1} mice are born in expected Mendelian ratios, are viable, fertile, and exhibit no obvious external limb phenotypes (Table 7). To assess potential limb skeletal phenotypes that were not immediately available upon gross evaluation of the external limbs I stained newborn *M1442*^{tm1} -/- pups with alizarin red and alcian blue. Alizarin red and alcian blue stained newborn *M1442*^{tm1} -/- pups were indistinguishable from wild-type littermates (Fig. 17). In particular, relative bone position and morphology of fore- and hindlimbs dissected away from alizarin red and alcian blue stained *M1442*^{tm1} -/- newborn pups appeared similar to wild-type limbs (Figs 18-19). Gross and histological sections of my *M1442*GWtg mice indicated that the development of digits I and V, and the posterior wrist and ankle elements may be adversely affected in *M1442*^{tm1} knockout mice; however, I noted no aberrations in these elements at this level of analysis (Figs. 18-19). Therefore, initial external and internal observations of limb morphology suggested that *M1442* is not required for proper limb development.

Table 7. Chi-square test for expected Mendelian ratios in offspring from heterozygous *M1442*^{tm1} intercrosses.

(N = 58)	WT	Het	Null
Expected:	14.5	29	14.5
Observed:	14	31	13

Two degrees of freedom
Chi-square = 0.32
Not significant at P = 0.5

Figure 17. Newborn and adult *M1442* null mice are indistinguishable from control littermates. (A-F) Control animals. (G-L) *M1442* null animals. (B, C, H, J) Dorsal and ventral forelimbs, respectively. (D, E, J, K) Dorsal and ventral hindlimbs, respectively. (F, L) Alizarin red and alcian blue stain for bone and cartilage in control (F) and mutant (L) newborn pups. Scale bars in (A-E, G-K) = 9 mm. Scale bars in (F, L) = 1 mm.

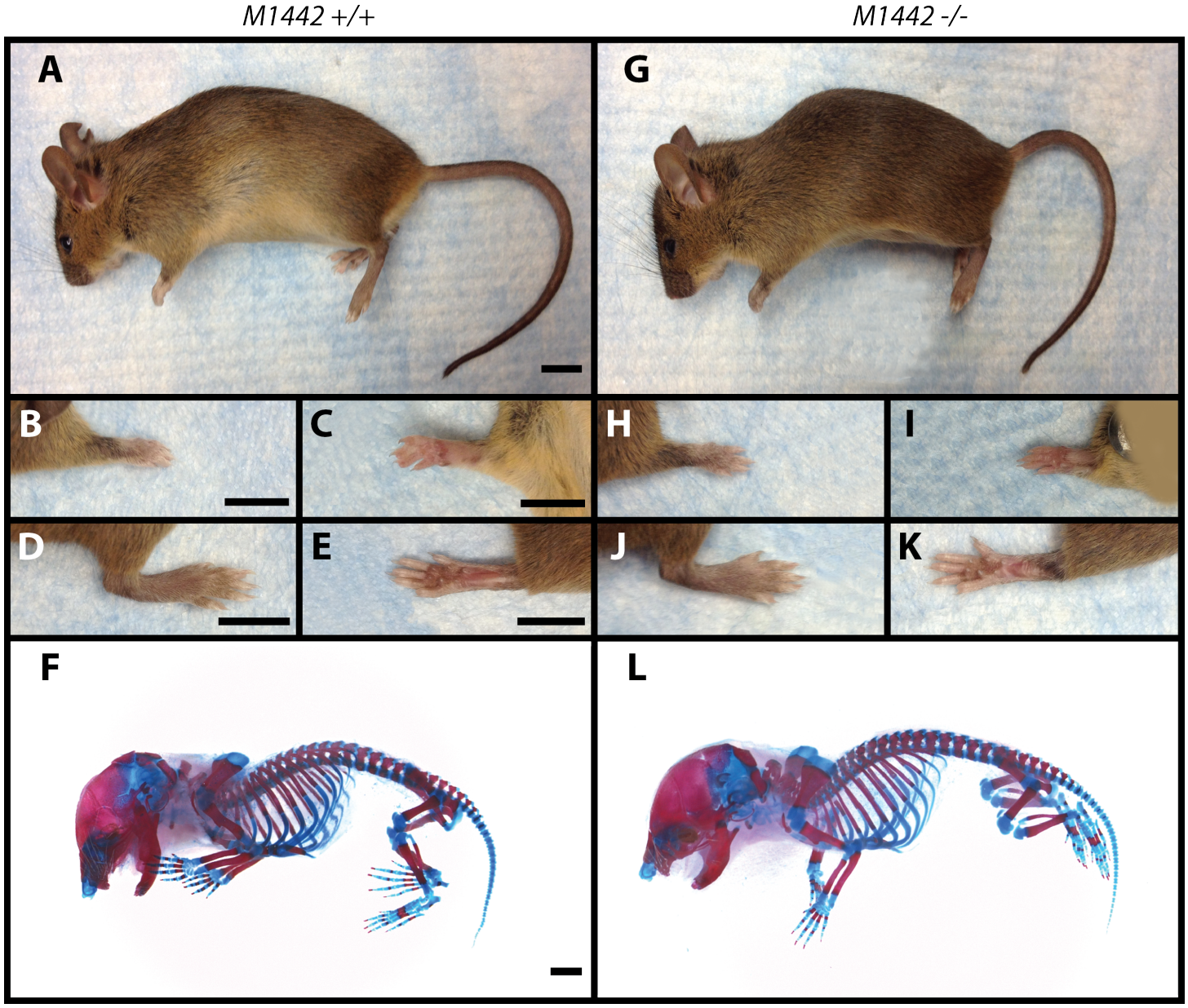

M1442 +/+
M1442 -/-
A
G
B
C
H
I
D
E
J
K
F
L

Figure 18. Forelimb autopod elements present in E18.5 and newborn *M1442* null embryos. (A-C) Control tissue. (D-F) *M1442* mutant tissue. (A, D) Alizarin red and alcian blue stain of newborn forelimbs. (B, E) H&E stained planar sections through forelimb autopods. (C, F) H&E stained frontal sections through forelimb autopods. Abbreviations: ca = capitate; dp = distal phalanx; ha = hamate; ip = intermediate phalanx; pp = proximal phalanx; ra = radius; tp = trapezoid; ul = ulna. Roman numerals correspond to digit rays.

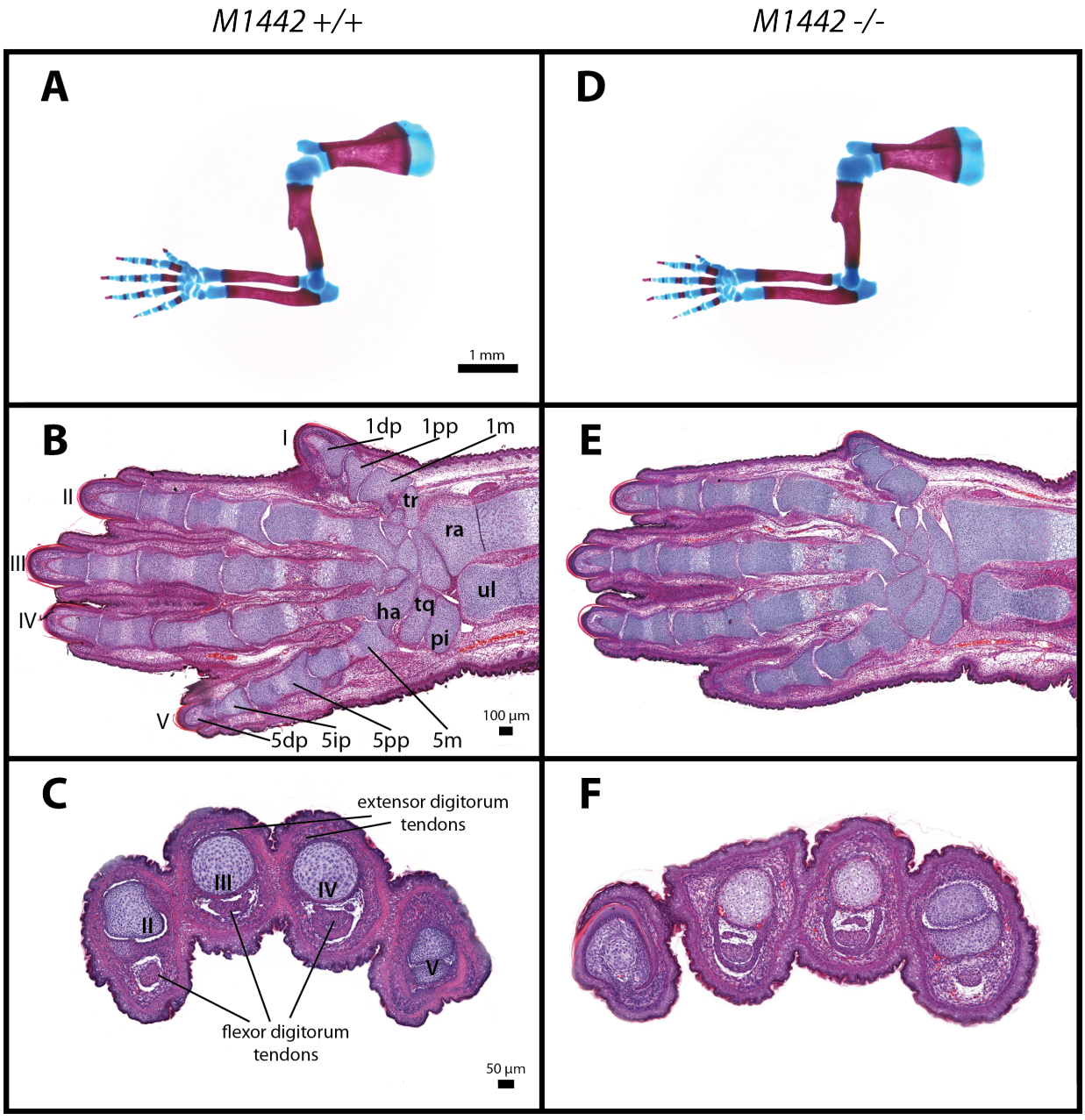
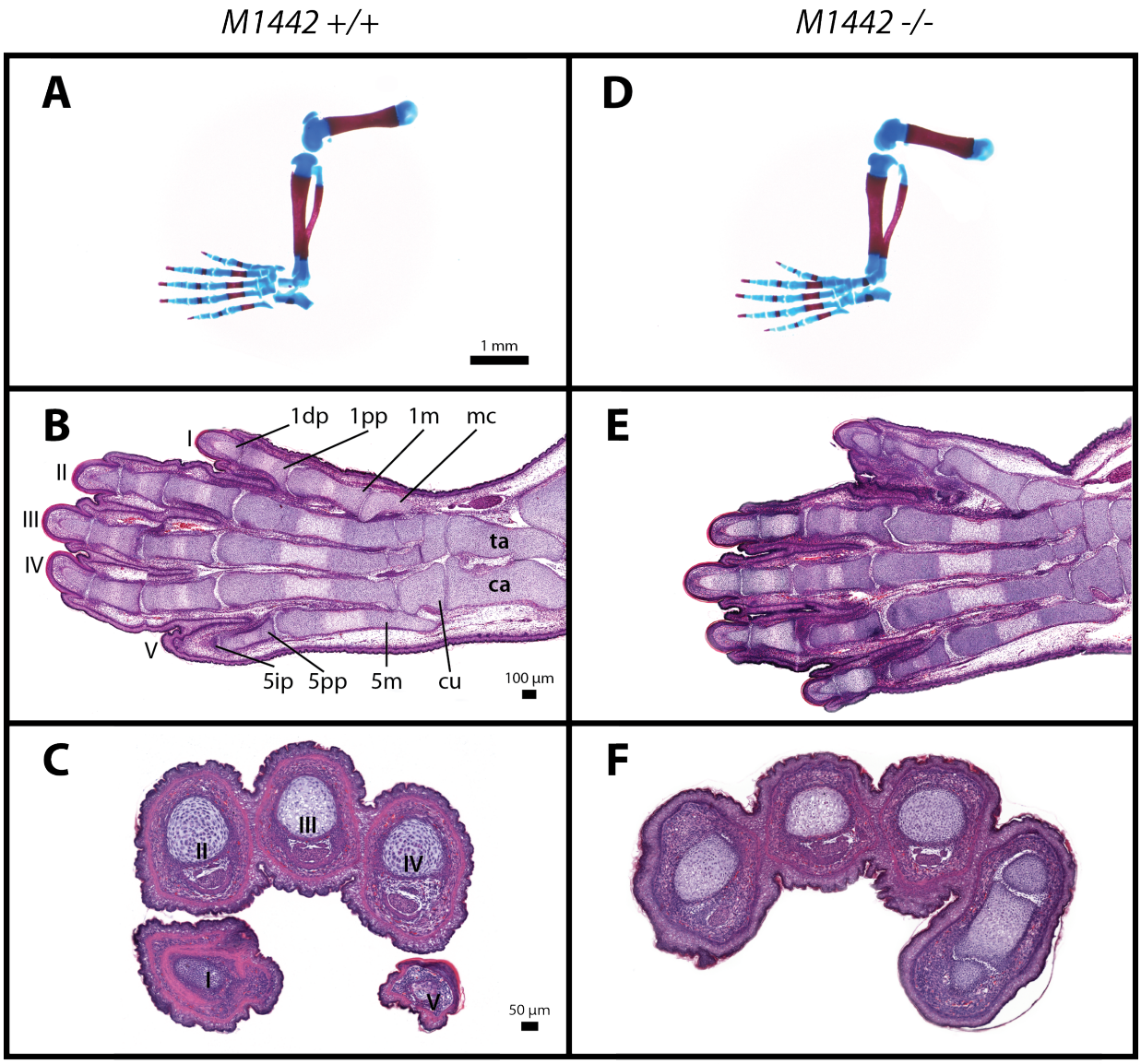

M1442 +/+
M1442 -/-
A
1 mm
D
B
I
1dp
1pp
1m
II
tr
ra
III
ul
ha
tq
IV
pi
V
5dp
5ip
5pp
5m
100 µm
E
C
extensor digitorum
tendons
III
IV
II
V
flexor digitorum
tendons
50 µm
F

Figure 19. Hindlimb autopod elements present in E18.5 and newborn *M1442* null embryos. (A-C) Control tissue. (D-F) *M1442* mutant tissue. (A, D) Alizarin red and alcian blue stain of newborn forelimbs. (B, E) H&E stained planar sections through hindlimb autopods. (C, F) H&E stained frontal sections through hindlimb autopods. Abbreviations: ca = calcaneus; cu = cuboid; dp = distal phalanx; ic = intermediate cuneiform; ip = intermediate phalanx; m = metatarsal; na = navicular; pp = proximal phalanx; ta = talus. Roman numerals correspond to digit rays.


M1442 +/+
M1442 -/-
A
1 mm
D
B
1dp
1pp
1m
mc
I
II
III
IV
V
ta
ca
5ip
5pp
5m
cu
100 µm
E
C
II
III
IV
I
V
50 µm
F

Quantitative analysis of transcription in enhancer-deleted mice

Though no gross anatomical aberrations accompanied the deletion of either *M1442* or *M280* I hypothesized that gene transcription was altered at the mutated loci. Both *M1442* and *M280* possess transcriptional enhancer activity, as demonstrated in our transgenic mice, where consistent and discreet enhancer-driven *lacZ* expression was noted in the developing limb buds. Moreover, *M1442* is located in the 19th intron of *Kifap3,* while *M280* lies within the second intron of *Rnf220*. It is possible that both enhancers regulate the limb-specific transcription of their host gene, a regulatory scenario documented before in the developing limb (Zhang and Williams, 2003; Donner and Williams, 2006). On the other hand, it is possible the enhancers regulate one or several neighboring genes, a regulatory scenario also documented previously (Zuniga et al., 2004).

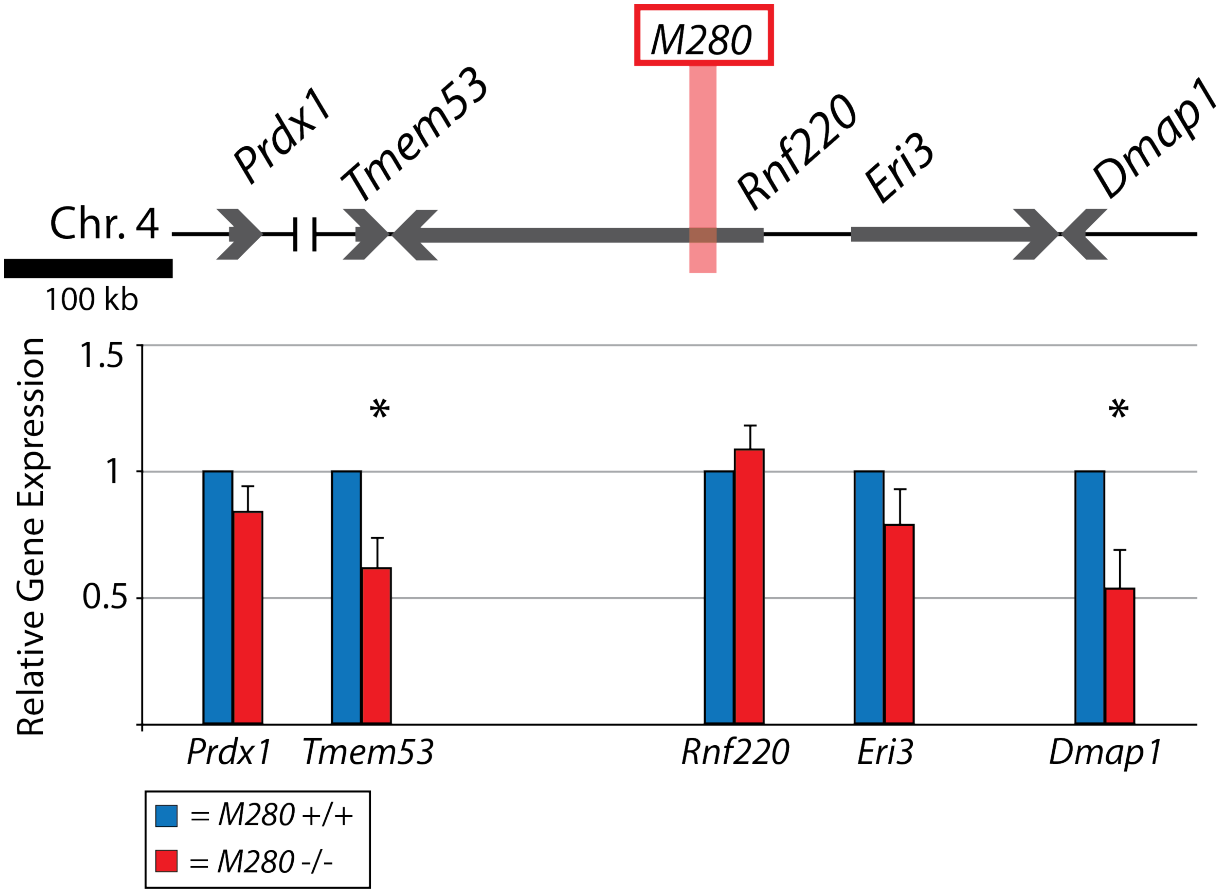
Because our selection of *M1442* and *M280* was not based on knowledge of the genes they regulate I created a logical framework to determine which genes to assay for transcriptional differences between mutant and wild-type animals: Briefly, I decided to assay the host gene, one or several genes immediately flanking the 5'- and 3'-margins of the host gene, additional genes beyond the immediate neighbors of the host gene if they appeared in a Pubmed and/or Google Scholar search in association with the key words "limb" and "development," and genes that were found up to 1 Mb upstream and downstream of the enhancers that were both 1) expressed at a statistically significant level in limb buds relative to other cell types (Cotney et al., 2012), and 2) associated with the key words "limb" and "development" in a Pubmed and/or Google Scholar search. The latter criterion stems from the observation that enhancers contributing to the transcriptional regulation of *Shh* and the *HoxD* gene cluster are located at distances approaching 1 Mb (Lettice et al., 2002; Lettice et al., 2003; Sagai et al., 2004; Montavon et al., 2011).

To investigate whether deletion of *M1442* and *M280* affected gene transcription I performed quantitative real-time PCR (qPCR) on mutant and wild-type E11.5 limb buds. Upon deletion of *M280* the transcription of its host gene, *Rnf220*, was unaffected in E11.5 limb buds, but two neighboring genes, *Tmem53* and *Dmap1*, experienced reduced transcription levels (Fig. 20; Appendix 1). The transcription level of *Tmem53* was just 60% of wild-type, while *Dmap1* was just 50% of wild-type. Transcription of another gene, *Eri3*, located in between *M280* and *Dmap1* was also unaffected, again demonstrating that an enhancer, *M280* in this case, can regulate multiple genes within a given window in what appears to be a selective manner.

Prxd1 is located 700 kb 5' of *M280*'s center-point, but previous research had indicated a role for this gene in interdigital cell death (Shan et al., 2005), a developmental phenomena that most certainly occurs in the spatial domain that overlaps with the *lacZ* expression pattern observed in our *M280-lacZ* transgenic mice (at E11.5 *lacZ* is observed in the autopod region that will give rise to digits II-IV). And per the criteria described above, *Prdx1* is upregulated in E10.5/E11.5 limb buds relative to two other cell types(Cotney et al., 2012). Therefore, I predicted that *M280* might regulate *Prdx1* transcription. However, absence of *M280* does not affect *Prdx1* transcription levels (Fig. 20).

I employed whole mount *in situ* hybridization (WMISH) to determine if *M280*'s host gene, *Rnf220*, is expressed in a limb-specific manner that overlaps with *M280* regulation (see Figs. 4, 5). Previous assays demonstrated that *Rnf220* exhibits a broad expression pattern, but the results were inconclusive as to whether it has a restricted limb expression pattern (www.informatics.jax.org; *Rnf220* primer ID: MTF1124). Interestingly, targeted deletion of the gene for the chromatin modifying protein, Sin3b, which interacts with the Rnf220-E3 ubiquitin ligase for its degradation, results in limb-specific shortening (David et al., 2008; Kong et al., 2010). Thus even though my qPCR results do not show a lessening of *Rnf220* levels, it is possible that additional enhancers compensate for *M280*'s absence and still regulate a central

Figure 20. Proximal-*M280* transcript levels in E11.5 limb buds. Assayed genes represented by gray arrows. The location of *M280* represented by vertical, red rectangle. Statistically significant transcript level changes at P = 0.05 represented by asterisk. (N=8 biological replicates for each genotype; see Materials and Methods).


M280
Prdx1
Tmem53
Rnf220
Eri3
Dmap1
Chr. 4
100 kb
Relative Gene Expression
1.5
1
0.5
*
*
Prdx1
Tmem53
Rnf220
Eri3
Dmap1
= M280 +/+
= M280 -/-

limb-specific *Rnf220* expression pattern. Also, expression of *Dmap1* was also assessed using WMISH because of the genes assayed with qPCR it showed the greatest change upon *M280* deletion.

WMISH revealed no observable differences in *Rnf220* or *Dmap1* expression pattern between E11.5 *M280* null and control embryos (Fig. 21). Indeed both genes did not appear to be expressed in a restricted manner whatsoever in the limb buds. The WMISH results suggest that *Rnf220* is not regulated in a strongly delimited manner in the limb, at least with the riboprobe used, and supports the conclusion from the qPCR results that *M280* does not regulated *Rnf220.* These results do not mean, however, that *M280* is not needed for *Dmap1* expression; rather, the resolution provided by WMISH may not be fine enough to detect transcript level changes of *Dmap1* in the limb.

Interestingly, deletion of *M1442* altered the transcription of numerous genes located approximately 500 kb 5' or 3' of *M1442* (Fig. 22; Appendix 1). Surprisingly, transcription of the host gene, *Kifap3*, increased by 30%, as did two neighboring genes, *Gorab* and *Scyl3*, which are located roughly 500 kb 5' and 30 kb 3' from the center-point of *M1442*, respectively. This result is unexpected because the transgenic mouse assays described earlier clearly indicate that *M1442* possess pro-transcriptional properties. The qPCR results suggest that manipulation of *M1442* may also disrupt transcriptional repression in a region that spans at least from *Gorab* to *Scyl3*. It is worth noting that the transcriptional output of *Kifap3*, *Gorab* and *Scyl3* was affected in the same way, an increase in expression of about 30%, suggesting that whatever repressive mechanism was disrupted by the removal of *M1442* uniformly affected this genomic window.

In contrast to the three previously discussed genes, the expression level of *Nme7*, which is located nearly 410 kb 3' of the *M1442* center-point, is down to 65% of wild-type. I assayed this gene for reasons other than the general framework I designed as a guide for selecting the others. It is clear that both *M1442* (transgenic mouse assays herein) and its human ortholog,

Figure 21. Whole mount *in situ* hybridization in control and M280 null E11.5 limb buds. Between eight and ten individuals were assessed per riboprobe.

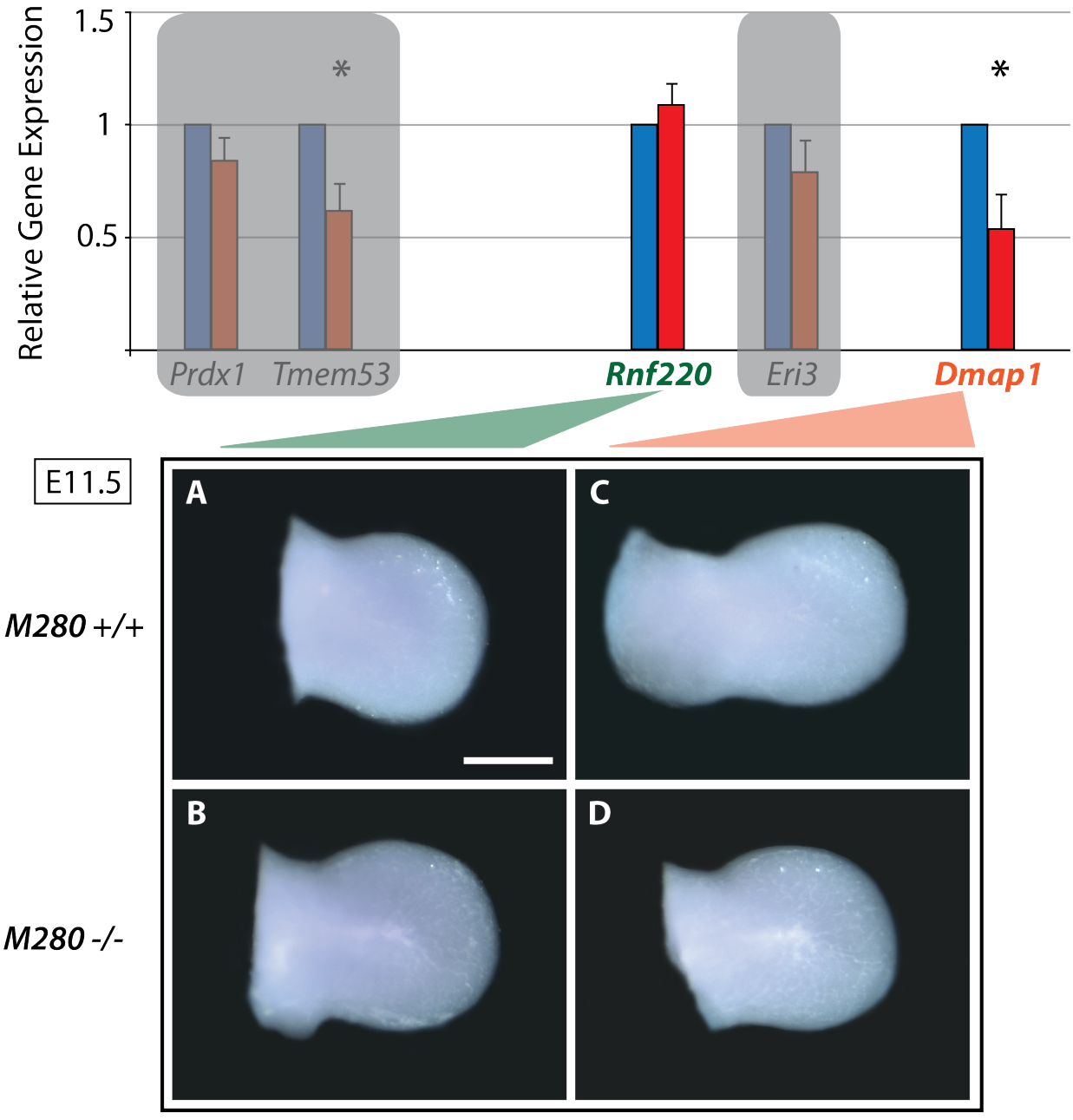
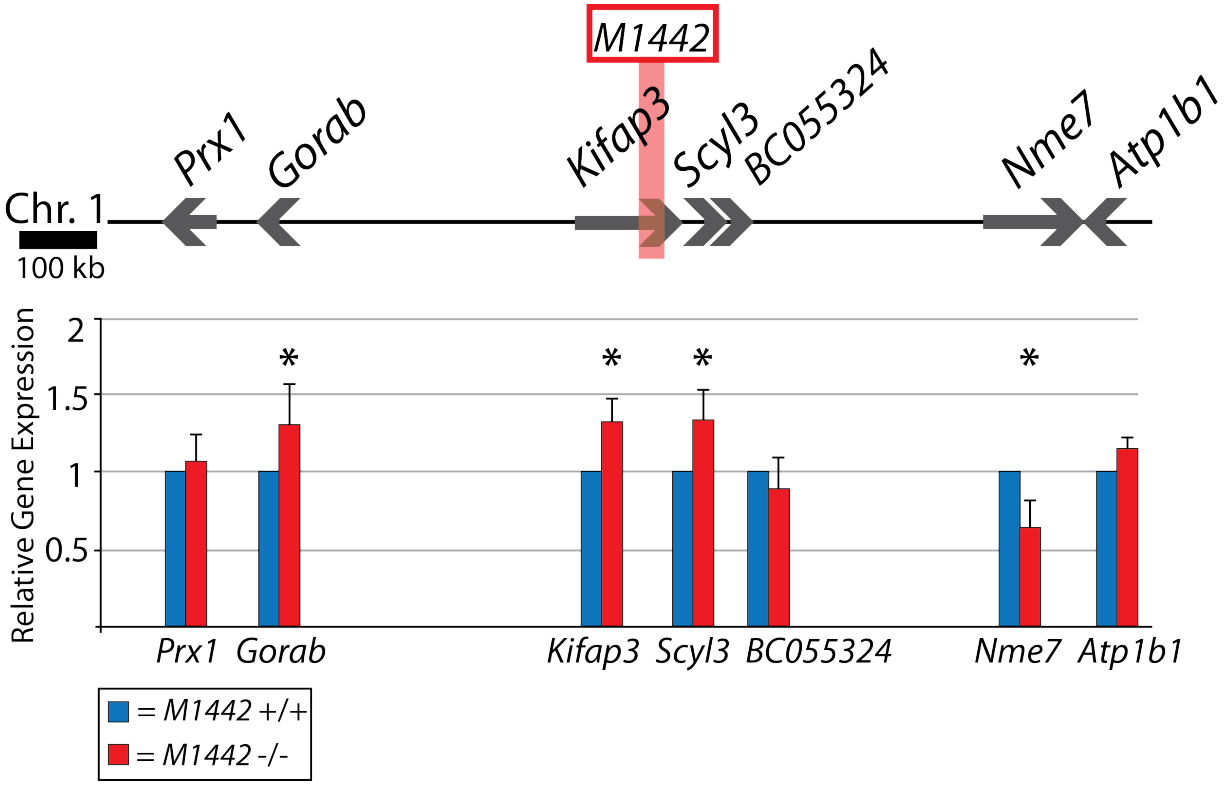

Relative Gene Expression
1.5
1
0.5
*
*
Prdx1
Tmem53
Rnf220
Eri3
Dmap1
E11.5
A
C
B
D
M280 +/+
M280 -/-

Figure 22. Proximal-*M1442* transcript levels in E11.5 limb buds. Assayed genes represented by gray arrows. The location of *M1442* represented by vertical, red rectangle. Statistically significant transcript level changes at P = 0.05 represented by asterisk. (N=8 biological replicates for each genotype; see Materials and Methods).


M1442
Prx1
Gorab
Kifap3
Scyl3
BC055324
Nme7
Atp1b1
Chr. 1
100 kb
Relative Gene Expression
2
1.5
1
0.5
*
*
*
*
Prx1
Gorab
Kifap3
Scyl3
BC055324
Nme7
Atp1b1
= M1442 +/+
= M1442 -/-

M1442, drive expression in the posterior margins of the E11.5 limb bud, which is also the site of required *Shh* expression (Wieczorek et al.; Riddle et al., 1993; Furniss et al., 2008) during limb development. Therefore, I surmised that perhaps genes associated with *Shh* signaling might be regulated by *M1442* because they too would presumably need to be expressed in the posterior E11.5 limb bud. In numerous developmental contexts *Shh* signaling requires cilia (Haycraft et al., 2005; Caspary et al., 2007; Goetz et al., 2009; Wong et al., 2009), as has been well documented, for example, in the establishment of cilia-dependent left-right asymmetry during vertebrate embryogenesis (Hamada et al., 2002). It is also known that genes in the Nme family are important for cilia function, and *Nme7* -/- mice display *situs inversus*, a failure to establish proper left-right asymmetry that leads to transposition of internal organs and vasculature (Boissan et al., 2009; Vogel et al., 2010). Therefore, the connection between the *Shh*-like expression pattern driven by *M1442* and the role of *Nme* genes in cilia motion made by the dependence of *Shh* signaling on cilia led me to predict that *M1442* was regulating *Nme7* transcription. It is worth noting that the transcription of two genes involved in ciliogenesis and movement, *Kifap3* and *Nme7*, are impacted by the absence of *M1442* in a statistically significant manner and that *M1442* directs transcription in a region of the developing limb that overlaps with *Shh* expression.

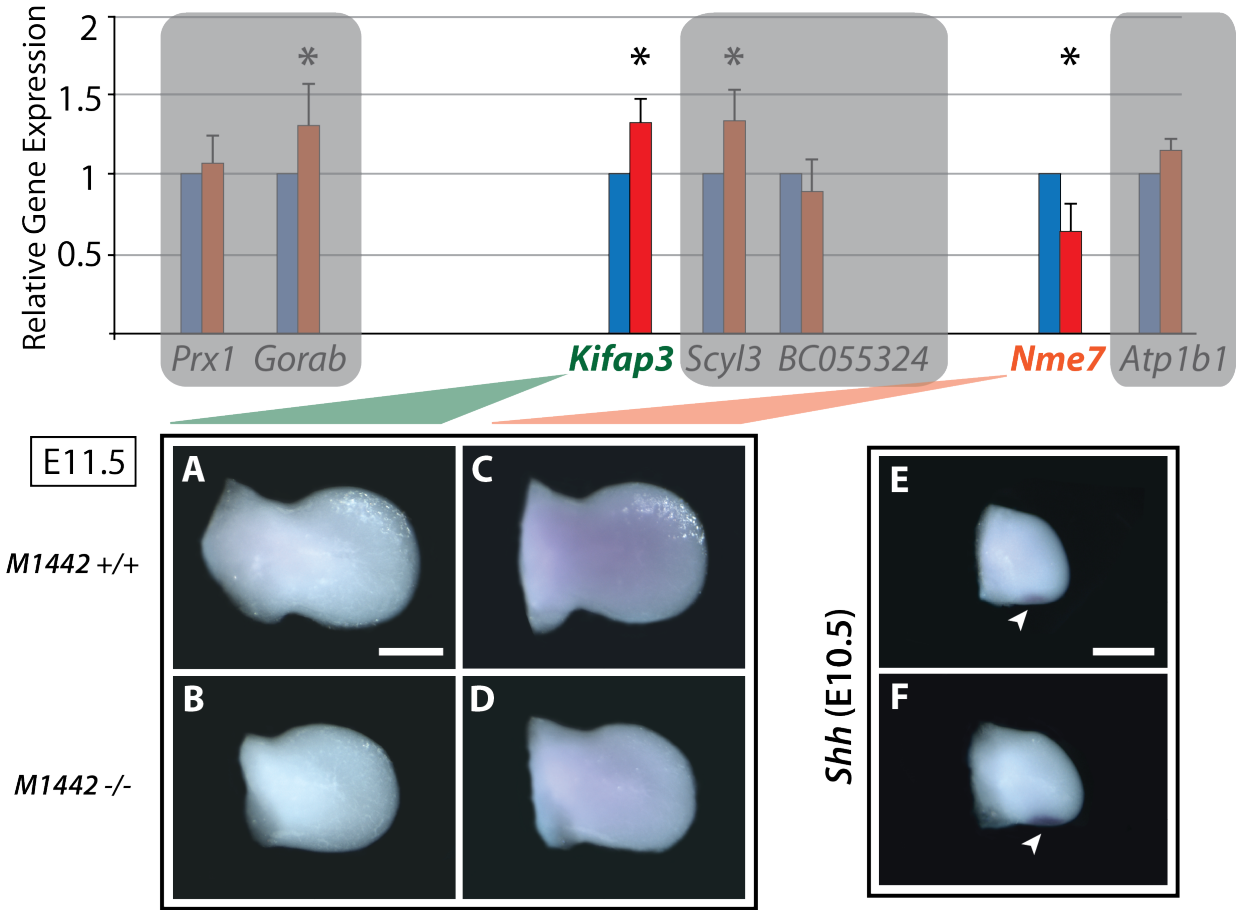
I also assayed the transcription of *Prx1*, a well-known transcription factor required for limb development (Martin et al., 1995), which is located approximately 600 kb 5' of *M1442*, and just 70 kb 5' of *Gorab*. Removal of *M1442* had no effect on *Prx1* transcription indicating that it does not regulate this gene despite the fact that the *Prx1* expression pattern overlaps with that observed in our *M1442-lacZ* transgenic mice. Moreover, this result reinforces the observation that for the genomic window assayed, the regulation encoded within *M1442* or extant because of *M1442*'s topographical location does not extend beyond *Gorab* in the 5' direction.

The transcription level of two other genes, *BC055324* and *Atp1b1*, were also unaffected by *M1442* deletion. Though *Atp1b1* did not meet all of the criteria above for selecting genes for qPCR – for example, *Atp1b1* is not associated with search terms "limb" and/or "development" in Google Scholar or Pubmed – but is was significantly up-regulated in E10.5 and E11.5 limb buds relative to two other cell types (Cotney et al., 2012) and I was curious to see if the positive transcriptional regulation conferred on *Nme7* by *M1442* extended to *Nme7*'s neighboring gene. These results hint at the complexity of the regulatory behavior of *M1442* because genes whose transcription is partially regulated by *M1442* are not necessarily clustered together. While *Kifap3*, *Gorab* and *Scyl3* are neighboring genes whose transcription is negatively regulated by *M1442*, *Nme7* is positively regulated by *M1442* even though the transcription of two assayed genes, *BC055324* and *Atp1b1*, which lie on either side of *Nme7* are unaffected.

I used WMISH to determine if *M1442*'s host gene, *Kifap3*, is expressed in a limb-specific manner that overlaps with *M1442* regulation (see Figs. 7, 9). *Kifap3* encodes the non-ATP using component of a kinesin motor. Such cellular machinery is required for ciliary mediated *Shh* expression (Kolpakova-Hart et al., 2007; Wong et al., 2009). Since the qPCR results indicated minor *Kifap3* expression changes, and since *M1442* directs expression in a region that overlaps with early limb bud *Shh* expression I wanted to see if WMISH could resolve a limb-specific *Kifap3* expression pattern. Expression of *Nme7* was also assessed using WMISH because of the genes assayed with qPCR it showed the greatest change upon *M1442* deletion. Additionally, a riboprobe for *Shh* allowed me to monitor its expression – this was important both as a control (the probe is very reliable) and because it would be curious to see if *Shh* expression is affected at all by *M1442* deletion as their domains of activity clearly overlap.

WMISH revealed no observable differences in *Kifap3* expression between E11.5 *M1442* null and control embryos (Fig. 23). Indeed *Kifap3* did not appear to be expressed in a restricted manner whatsoever in the limb buds. *Nme7* does not appear to be expressed in a spatially

Figure 23. Whole mount *in situ* hybridization in control and *M1442* null E11.5 limb buds. Between eight and ten individuals were assessed per riboprobe.

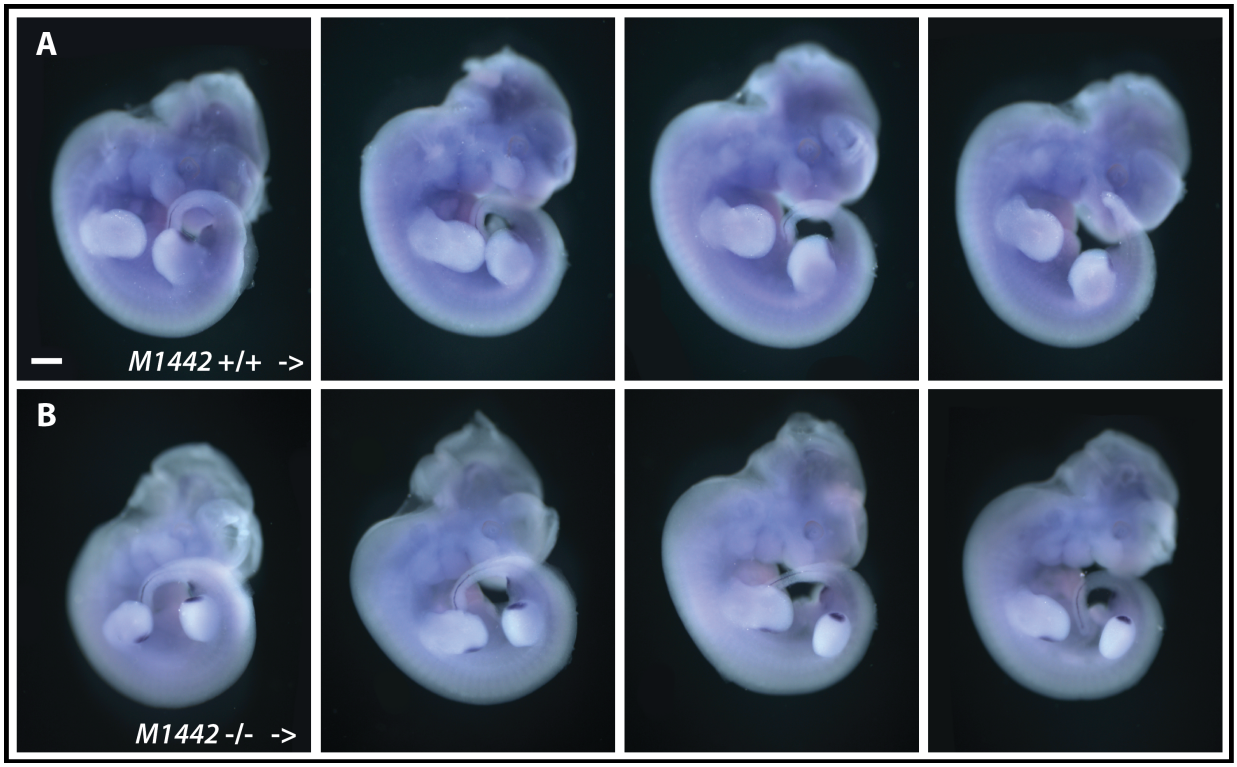

Relative Gene Expression
2
1.5
1
0.5
*
*
*
*
Prx1
Gorab
Kifap3
Scyl3
BC055324
Nme7
Atp1b1
E11.5
M1442 +/+
M1442 -/-
A
B
C
D
E
F
Shh (E10.5)

restricted manner in the E11.5 limb bud; however, it is possible to discern a slight decrease in signal intensity in the mutant bud. The WMISH results suggest that *Kifap3* is not regulated in a strongly delimited manner in the limb, at least with the riboprobe used. These results do not mean, however, that *M1442* does not impinge on *Kifap3* expression as indicated by qPCR; rather, the resolution provided by WMISH may not be fine enough to detect transcript level changes of *Kifap3* in the limb. On the other hand, these results indicate that the *M1442* regulation of *Kifap3* indicated by qPCR is not likely to be a required input for proper limb development. Interestingly, although classic *Shh*-related polydactyly/limb patterning phenotypes accompany *M1442*'s deletion, *Shh* expression appeared to me to be consistently more intense in E10.5-11.5 *M1442* null embryos relative to controls (Fig. 24).

Comparative dissection and characterization of mutant mouse limbs

To this point, it has been demonstrated that gross anatomical limb aberrations did not accompany deletion of either *M280* or *M1442;* nor did histological sections reveal any absences or deficiencies in bone elements and growth. Furthermore, qPCR showed minor transcription changes in mutant animals but no obvious connections could be made between the enhancers and their regulatory targets. Inasmuch as a lack of a phenotype is more difficult to demonstrate than the of presence of one, I found it necessary to employ more traditional approaches to phenotypic analysis to show that mutant autopod elements were indeed indistinguishable from controls in shape and location: dissection, diagrammatic recording and comparison were in order. More specifically, forelimbs extracted from at least three E15.5 and 18.5 embryos were analyzed carefully after alcian blue (both E15.5 and 18.5) and alizarin red (E18.5 only) staining from each genotype: control littermates, $M280^{tm1}$ -/-, and $M1442^{tm1}$ -/-. Under the microscope I looked at relevant autopod elements from each embryonic limb and recorded their presence,

Figure 24. Limb-specific *Shh* expression more intense in E10.5-11.5 *M1442* null embryos than in controls.


A
M1442 +/+ ->
B
M1442 -/- ->
—

position and shape by making free-hand sketches. Though this approach is not as sophisticated as others I've described I think it gave me a very good, almost automatic, impression of the integrity or soundness of a particular embryo's autopod structure. Moreover, the positioning and shape of wrist elements are, for lack of a better word, a bit confusing (especially in cross-section), but by taking the time to analyze individual wrists and sketch their anatomical structures I became more confident that indeed my mutant mice, both *M280*^{tm1} -/- and *M1442*^{tm1} -/-, did not manifest any obvious morphological defects. In particular, digits II-IV and the wrist elements just proximal to them – the trapezoid, capitate and hamate – were indistinguishable between *M280*^{tm1} -/- and control limbs (Fig. 25). And in *M1442*^{tm1} -/- limbs, the elements of digits I and V, and the nearby wrist elements – the trapezium, hamate, triquetral and pisiform – were indistinguishable from control limbs (Figs. 26-28). Histological sections of E18.5 *M1442* forelimbs also confirmed that development of the carpal elements that experience *M1442*-regulated gene expression was unaffected (Fig. 29).

Fine-scale length analysis of autopod bone elements in mutant mice

I decided to look for subtle defects in the development of digits in my mutant mice. Subtle changes in digit bone length had been reported before in another limb-enhancer deletion experiment – the *Tbx4* limb enhancer. Normal hindlimb development requires early *Tbx4* expression (Hasson et al.; Naiche et al.; Naiche and Papaioannou, 2003). Although mice lacking a particular enhancer for *Tbx4* appeared to be grossly normal, with all hindlimb bones present (though some were modified slightly or fused), it was shown that individual bone lengths were altered. For example, digit ray I (metacarpal to distal phalanx) was reduced in size by 8%, while digit ray V was reduced by 3% (Menke et al., 2008).

Figure 25. E18.5 *M280* null autopods are anatomical indistinguishable from controls.

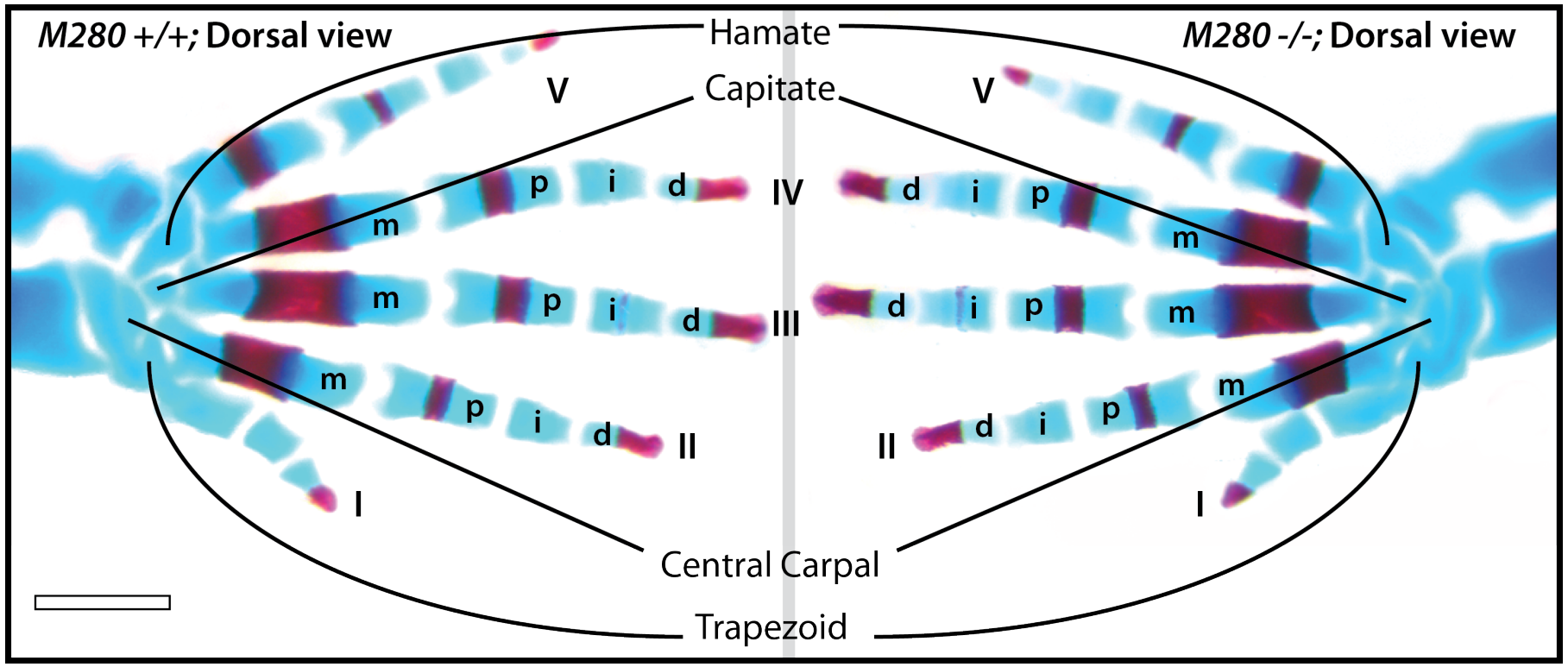
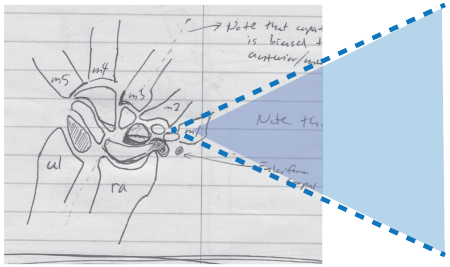
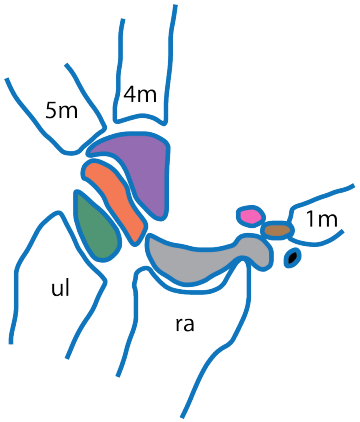

M280 +/+; Dorsal view
M280 -/-; Dorsal view
Hamate
Capitate
Central Carpal
Trapezoid
V
IV
III
II
I
m
p
i
d

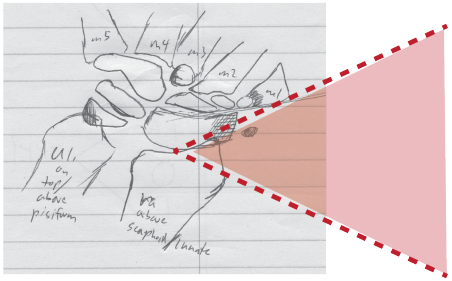
Figure 26. Dorsal view of E15.5 *M1442* null and control forelimb carpal regions.

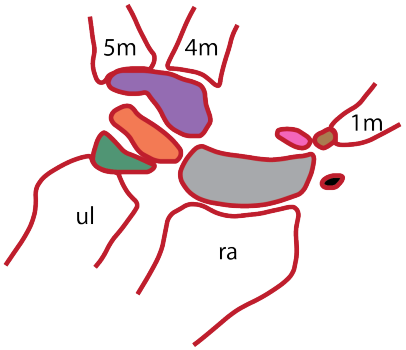
M1442 +/+; Dorsal view





M1442 -/-; Dorsal view





Hamate

Triquetral

Pisiform

Trapezium

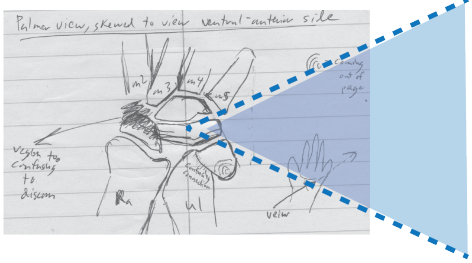
Trapezoid

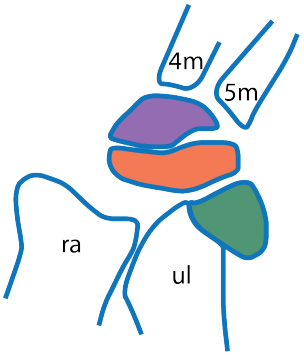
Scaphoid-Lunate

Falciform carpal

Figure 27. Palmer/ventral view of E15.5 *M1442* null and control forelimb carpal regions.

***M1442* +/+; Ventral-lateral view**



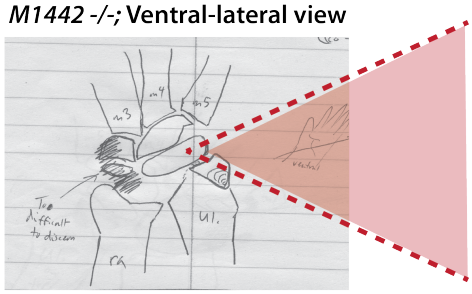


Hamate

Triquetral

Pisiform

***M1442* -/-; Ventral-lateral view**



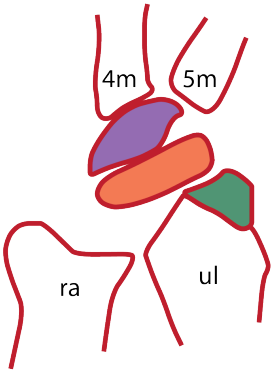
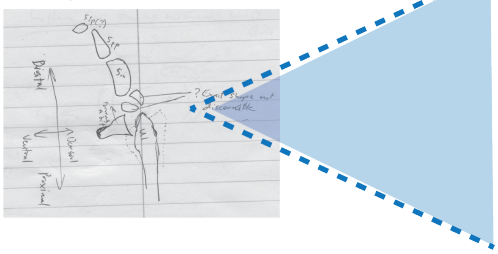
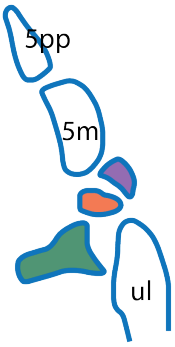


Figure 28. Left lateral view of E15.5 *M1442* null and control forelimb carpal regions.

M1442 +/+; Lateral-anterior view



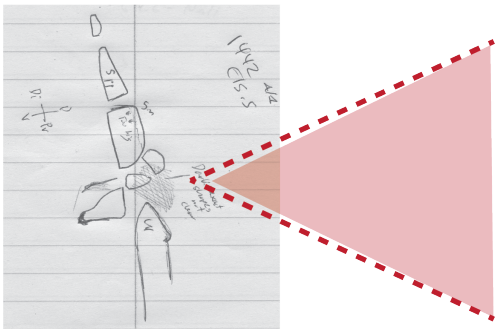


Hamate

Triquetral

Pisiform

M1442 -/-; Lateral-anterior view



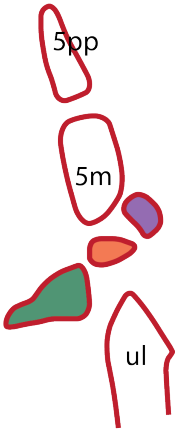
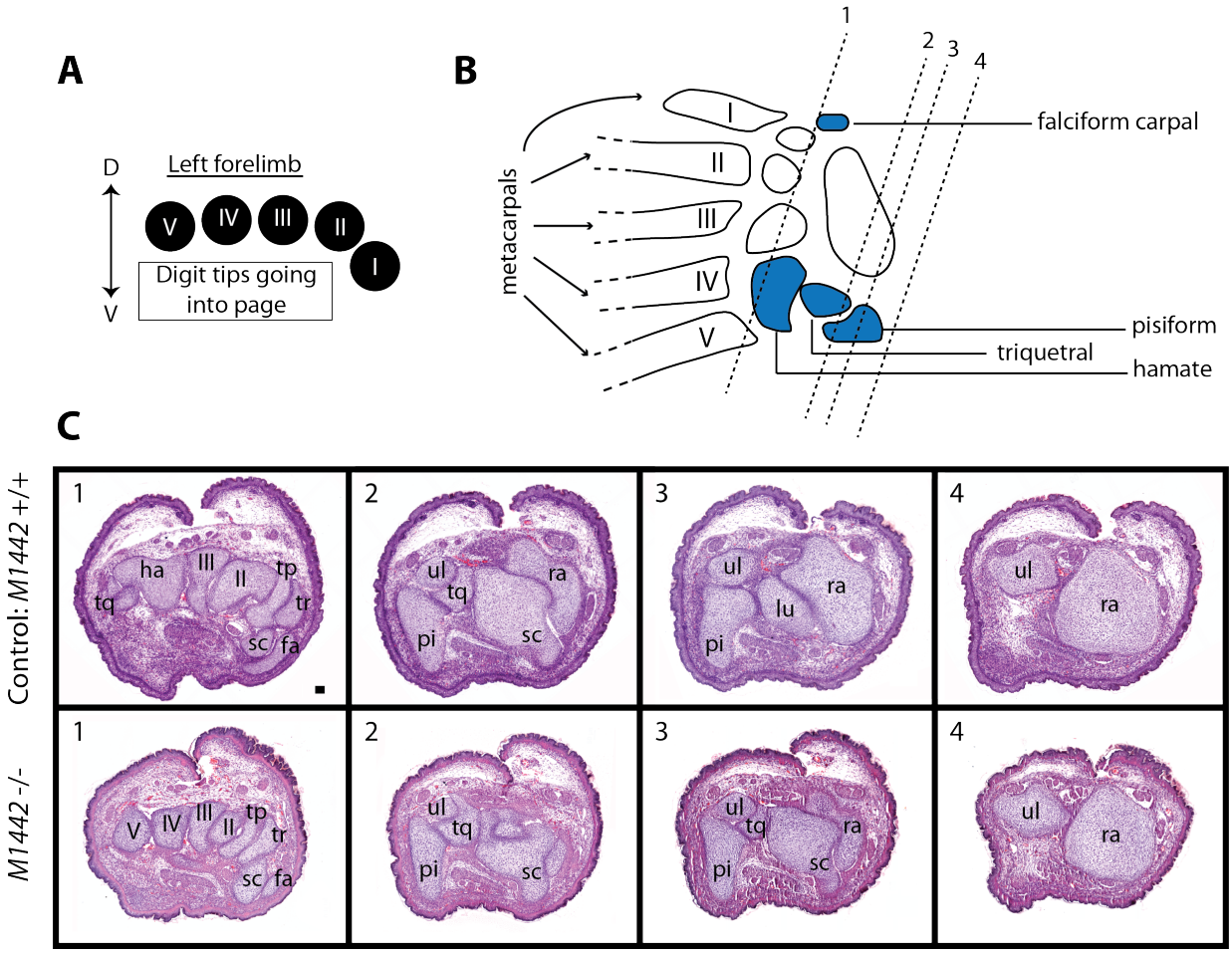


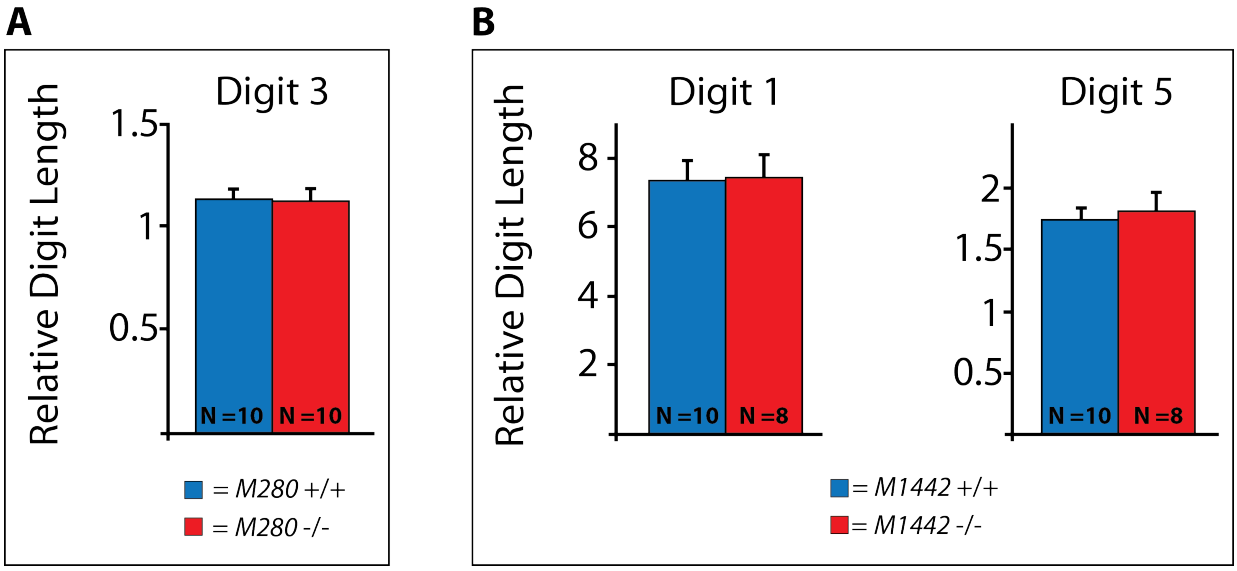
Figure 29. Cross-sections through E18.5 *M1442* null and control forelimb carpal regions. (A) Orientation diagram; D = dorsal, V = ventral. (B) Diagrammatic representation of sectioned forelimb carpal elements. Arabic numerals correspond to numerals in upper and lower rows in (C). Blue shading indicates elements positive for *lacZ* signal in *M1442*-GWtg lines. (C) H&E stained sections of E18.5 forelimb carpal regions. Roman numerals refer to digit rays. Abbreviations: fa = falciform carpal; pi = pisiform; ra = radius; sc = scaphoid; tp = trapezoid; tr = trapezium; ul = ulna. Scale bar is 50 μm.


A
D
Left forelimb
V
IV
III
II
I
Digit tips going into page
V
B
1
2
3
4
metacarpals
I
II
III
IV
V
falciform carpal
pisiform
triquetral
hamate
C
Control: M1442 +/+
M1442 -/-
ha
tq
tp
tr
sc
fa
ul
ra
pi
lu

To assess whether *M280*^{tm1} -/- mice manifested minor, but significant decreases in the length of autopod elements I collected ten mutant and control E18.5 embryos and dissected away the forelimbs. The forelimbs were then stained with alizarin red and alcian blue – dyes which stain the bones and cartilage, respectively. Weights of each sacrificed embryo were also recorded. Results from our *M280*GWtg experiments indicated that the central digit ray, digit ray III, would be more likely than other elements to undergo aberrant development in the absence of *M280*. I measured the length of each of the four individual elements of digit ray III – 3rd metacarpal, and the proximal, intermediate and distal phalanges – in each mutant and control animal in both right and left limbs. These measurements were then added together to achieve a total length for digit ray III and the lengths of digit III from both right and left limbs were averaged together. When digit III lengths were normalized to radii length (radius length reflects embryo size, recorded as embryonic weight; Pearson product-moment correlation coefficient, r = 0.54; Appendix 1) the length of digit III was not statistically different between *M280* null and control embryos (Fig. 30A). Therefore, I concluded that digit III is not dependent on the presence of *M280*, not even for minor inputs into its developmental integrity.

To assess whether *M1442*^{tm1} -/- mice manifested minor, but significant decreases in the length of autopod elements I collected eight mutant and ten control E18.5 embryos and dissected away the forelimbs. The forelimbs were then stained with alizarin red and alcian blue – dyes which stain the bones and cartilage, respectively. Weights of each sacrificed embryo were also recorded. Results from my *M1442*GWtg experiments indicated that the anterior-most and posterior-most digit rays, digit rays I and V, would be more likely than other elements to undergo aberrant development in the absence of *M1442*. In each mutant and control forelimb I measured the length of each of the three individual elements of digit ray I (1st metacarpal, and the proximal and distal phalanges) and each of the four individual elements of digit ray V (4th metacarpal, and the proximal, intermediate and distal phalanges). These measurements were

Figure 30. Digit lengths in mutant *M280* and *M1442* mice. (A) Digit III lengths normalized to radius lengths in E18.5 *M280* null and control embryos. (B) Digits I and IV lengths normalized to radius lengths in E18.5 *M1442* null and control embryos.


A
Digit 3
Relative Digit Length
1.5
1
0.5
N =10
N =10
= M280 +/+
= M280 -/-
B
Digit 1
Digit 5
Relative Digit Length
8
6
4
2
2
1.5
1
0.5
N =10
N =8
N =10
N =8
= M1442 +/+
= M1442 -/-

then added together to achieve a total length for digit I and V, respectively, and the lengths of either digits I or V from both right and left limbs were averaged together. When digit I lengths were normalized to radii length the length of digit I was not statistically different between *M1442* null and control embryos (Fig. 30B). The same was true for digit V (Fig.30B). Therefore, I concluded that digits I and V are not dependent on the presence of *M1442*, not even for minor inputs into their development along the proximal-distal axis of the autopod.

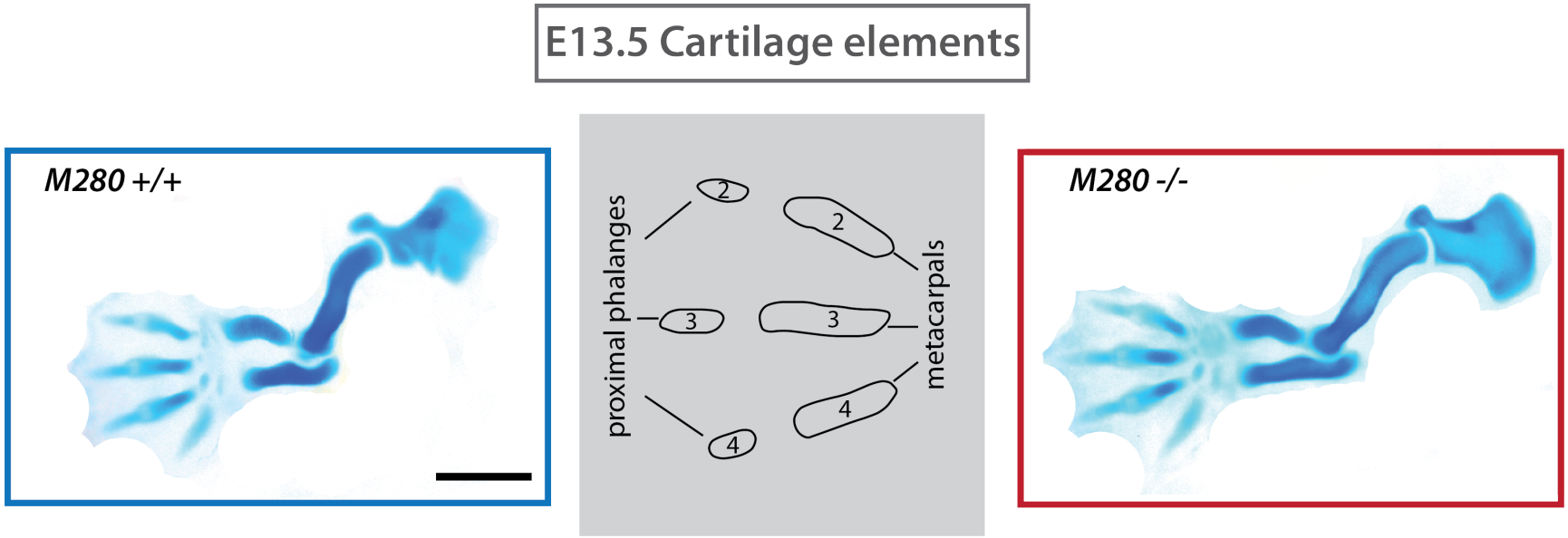
Developmental progression of cartilaginous limb elements in mutant mice

It is possible that the absence of *M280* in mutant embryos results in transient defects in limb development. To test this hypothesis I collected mutant and control embryos from at least three litters generated from $M280^{tm1}$ -/+ intercrosses at three embryonic time points: E13.5, E14.5, and E15.5. Embryos were then stained with alcian blue to visualize the developmental progression of cartilaginous limb elements. The lower time point, E13.5, was selected, because readily identifiable cartilaginous autopod elements become discernable at this time. By E15.5 each of the cartilaginous precursors of the digit bones are present.

In summary, I observed no indication that there is a transient developmental delay in the progression of autopod elements from E13.5 to E15.5 in *M280* null embryos (Figs. 31-33). However, at E15.5 more than half the mutant-control pairwise comparisons showed that mutant limbs were smaller than controls. In E13.5 limbs I assessed the sizeand shape of the metacarpals and proximal phalanges of digits II-IV. (Only two litters were assessed at E13.5; it seems that with this time point Mendelian odds were not on my side!). Of seven mutant-control pairwise comparisons at E13.5 five scored as equal and two scored as mutant-larger-than-control. In E14.5 limbs I assessed the size and shape of metacarpals and the proximal and intermediate phalanges in digits II-IV. Of five mutant-control pairwise comparisons at E14.5 three were

Figure 31. Forelimb comparisons between *M280* null and control embryos at E13.5. (A) Control tissue outlined in blue; mutant tissue outlined in red. Grey box shows cartilage elements used in comparison. (B) Each litter is represented by a different colored box: orange or green. Control limbs (blue) from each litter indicated by Arabic numerals. Mutant limbs (red) from each litter indicated by lowercase letters. If control limb is shorter (and/or more delayed) or larger (and/or more advanced) than mutant limb then it will be indicated by a dotted outline. Scale bars = 0.5 mm.

A


E13.5 Cartilage elements
M280 +/+
proximal phalanges
metacarpals
2
2
3
3
4
4
M280 -/-

B

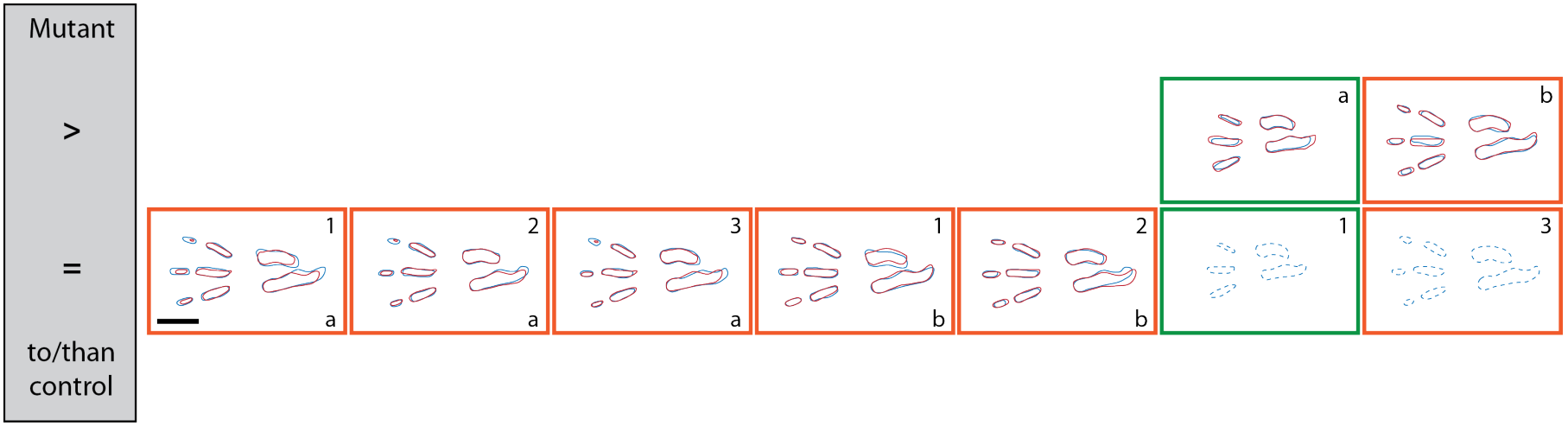

Mutant
>
=
to/than control
1 a
2 a
3 a
1 b
2 b
a
b
1
3

Figure 32. Forelimb comparisons between *M280* null and control embryos at E14.5.
(A) Control tissue outlined in blue; mutant tissue outlined in red. Grey box shows cartilage elements used in comparison. (B) Each litter is represented by a different colored box: orange, green or purple. Control limbs (blue) from each litter indicated by Arabic numerals. Mutant limbs (red) from each litter indicated by lowercase letters. If control limb is shorter (and/or more delayed) or larger (and/or more advanced) than mutant limb then it will be indicated by a dotted outline. Scale bars = 0.5 mm.

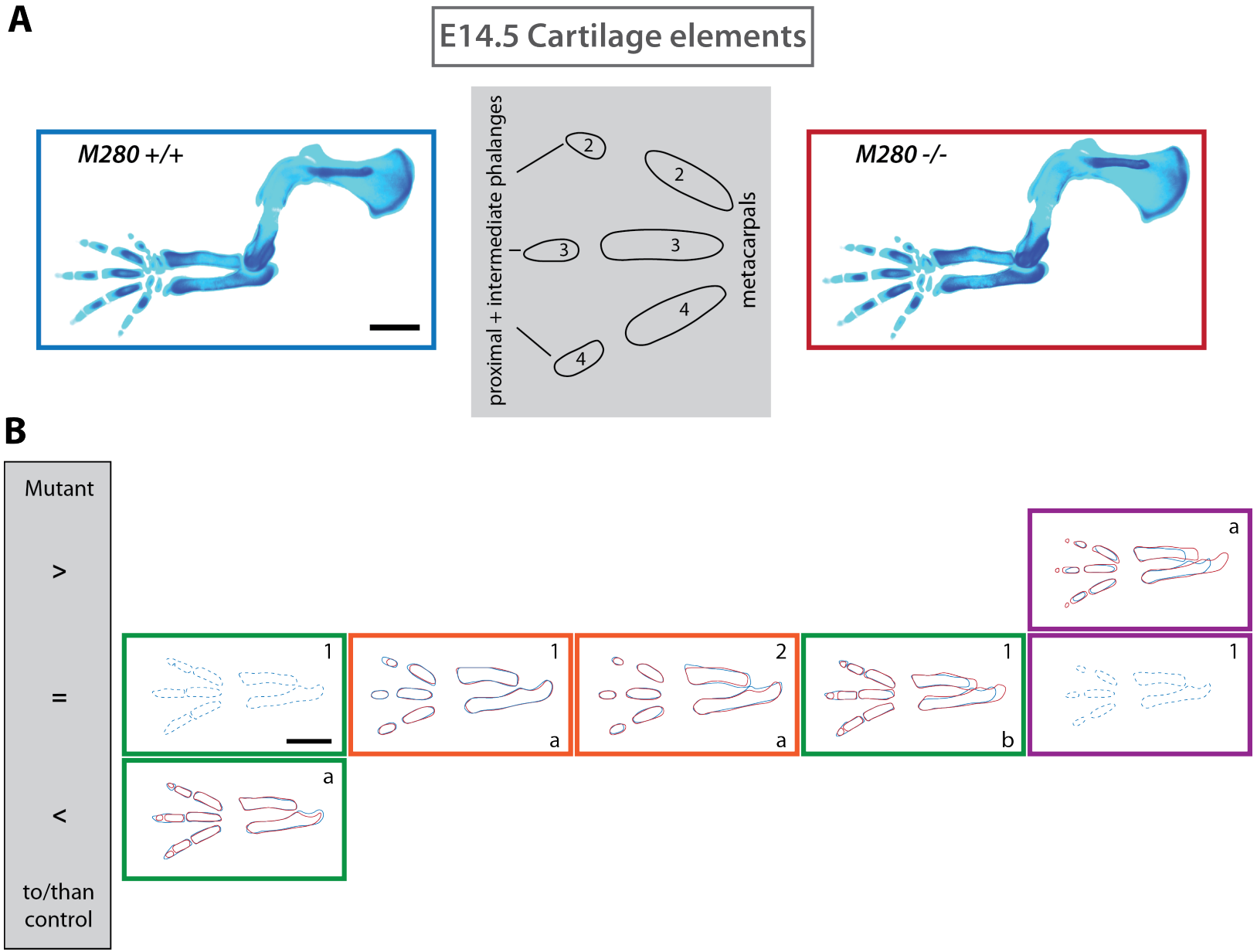
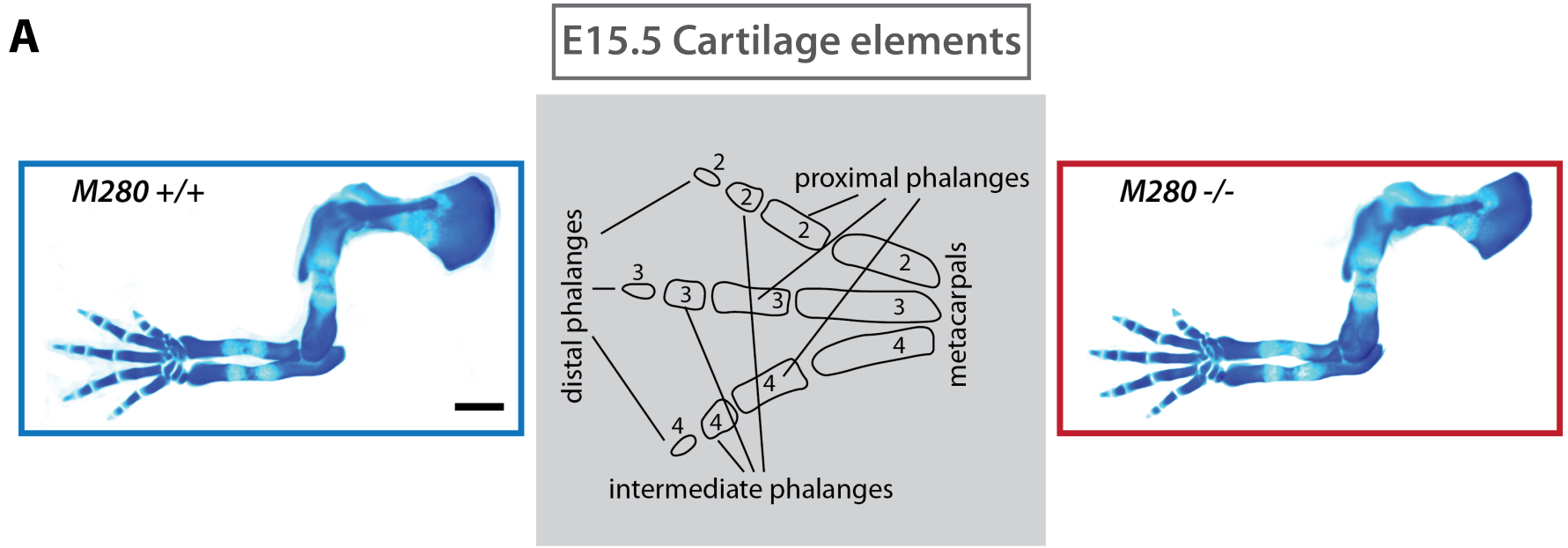

A
E14.5 Cartilage elements
M280 +/+
proximal + intermediate phalanges
metacarpals
2
2
3
3
4
4
M280 -/-
B
Mutant
>
=
<
to/than control
1
1
a
2
a
1
b
1
a
a

Figure 33. Forelimb comparisons between *M280* null and control embryos at E15.5.
(A) Control tissue outlined in blue; mutant tissue outlined in red. Grey box shows cartilage elements used in comparison. (B) Each litter is represented by a different colored box: orange, green, purple or brown. Control limbs (blue) from each litter indicated by Arabic numerals. Mutant limbs (red) from each litter indicated by lowercase letters. If control limb is shorter (and/or more delayed) or larger (and/or more advanced) than mutant limb then it will be indicated by a dotted outline. Scale bars = 0.5 mm.


A
E15.5 Cartilage elements
M280 +/+
M280 -/-
proximal phalanges
distal phalanges
metacarpals
intermediate phalanges
B
Mutant
=
<
to/than
control

scored as equal, one was scored as mutant-smaller-than-control, and one was scored as mutant-larger-than-control. In E15.5 limbs I assessed the size and shape of all four digit elements, from metacarpals to distal phalanges, in digits II-IV. Of nine mutant-control pairwise comparisons four were scored as equal and five were scored as mutant-smaller-than-control. The majority of all pairwise comparisons over the three time points were scored as equal, but as mentioned, at E15.5 I recorded a trend towards shorter mutant limbs.

I also pursued the hypothesis that *M1442*^{tm1} -/- embryos experienced a transient development defect in limb/autopod development. As was done with my *M280* null mice, I tested this hypothesis by collecting mutant and control embryos from at least three litters generated from *M1442*^{tm1} -/+ intercrosses at three embryonic time points: E13.5, E14.5, and E15.5. The assessment of potential developmental defects in limb progression was then carried out in the exact manner as that described above for *M280* null-control comparisons, with the notable exception that I focused on digits I and V, as these were the elements expressing *lacZ* in my *M1442*GWtg mice.

In summary, I observed no indication that there is a transient developmental defect in the progression of autopod elements from E13.5 to E15.5 in *M1442* null embryos (Figs. 34-36). In E13.5 limbs I assessed the size and shape of the 5th metacarpal and the hamate and triquetral, wrist bones near the 5th metacarpal. Of twelve mutant-control pairwise comparisons at E13.5 seven scored as equal and five scored as mutant-larger-than-control. In E14.5 limbs I assessed the size and shape of the first and fifth metacarpals, the proximal phalange of digit I, the proximal and intermediate phalanges of digit V, and the hamate and triquetral. Of nine mutant-control pairwise comparisons at E14.5 seven were scored as equal and two scored as mutant-smaller-than-control. In E15.5 limbs I assessed the size and shape of all four digit elements, from metacarpals to distal phalanges, in digits I and V, as well as the hamate, triquetral and pisiform bones of the wrist. Of six mutant-control pairwise comparisons four were scored as

Figure 34. Forelimb comparisons between *M1442* null and control embryos at E13.5.
(A) Control tissue outlined in blue; mutant tissue outlined in red. Grey box shows cartilage elements used in comparison. (B) Each litter is represented by a different colored box: orange, green or purple. Control limbs (blue) from each litter indicated by Arabic numerals. Mutant limbs (red) from each litter indicated by lowercase letters. If control limb is shorter (and/or more delayed) or larger (and/or more advanced) than mutant limb then it will be indicated by a dotted outline. Scale bars = 0.5 mm.

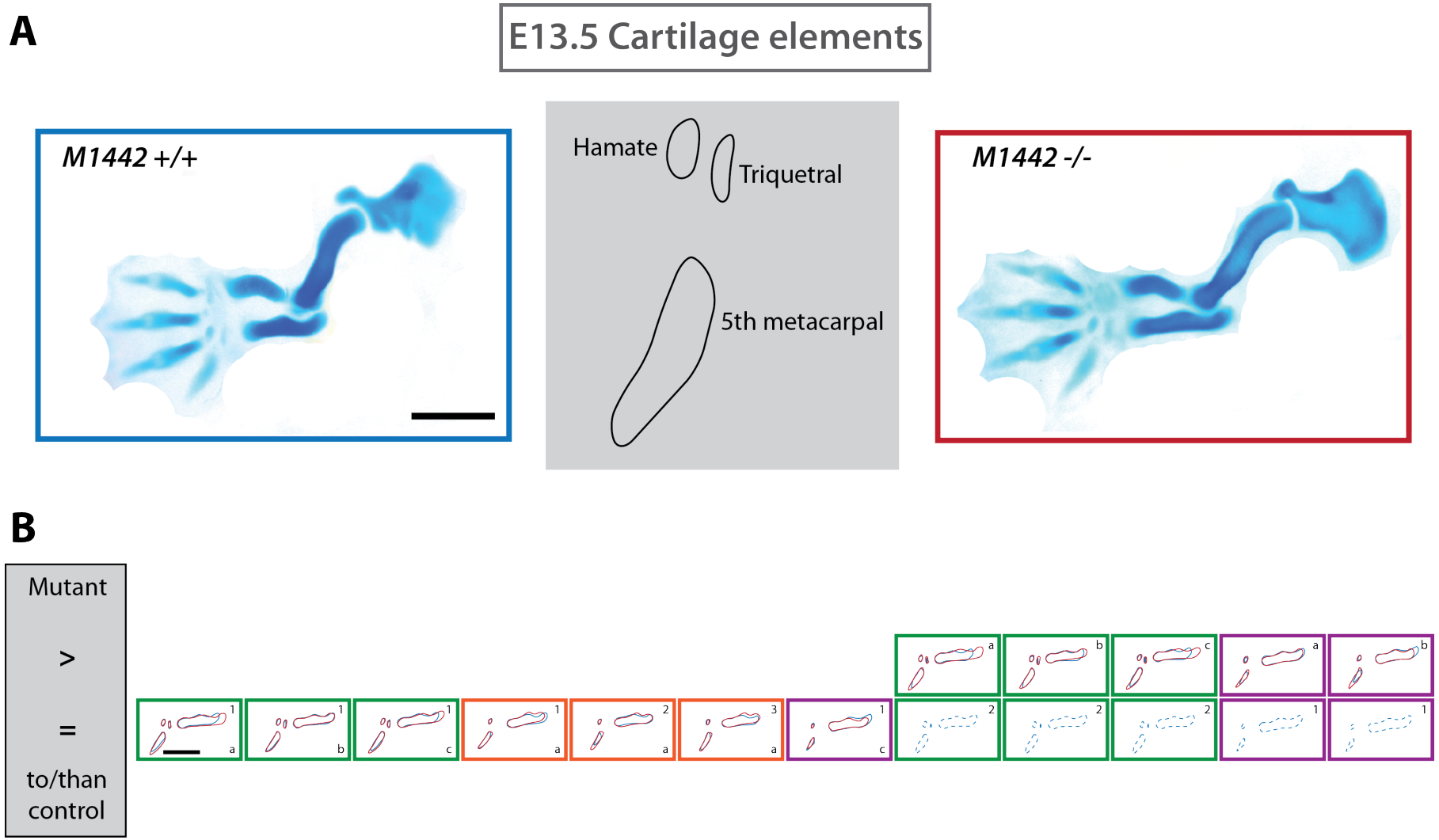

A
E13.5 Cartilage elements
M1442 +/+
Hamate
Triquetral
5th metacarpal
M1442 -/-
B
Mutant
>
=
to/than
control

Figure 35. Forelimb comparisons between *M1442* null and control embryos at E14.5.
(A) Control tissue outlined in blue; mutant tissue outlined in red. Grey box shows cartilage elements used in comparison. (B) Each litter is represented by a different colored box: orange, green or purple. Control limbs (blue) from each litter indicated by Arabic numerals. Mutant limbs (red) from each litter indicated by lowercase letters. If control limb is shorter (and/or more delayed) or larger (and/or more advanced) than mutant limb then it will be indicated by a dotted outline. Scale bars = 0.5 mm.

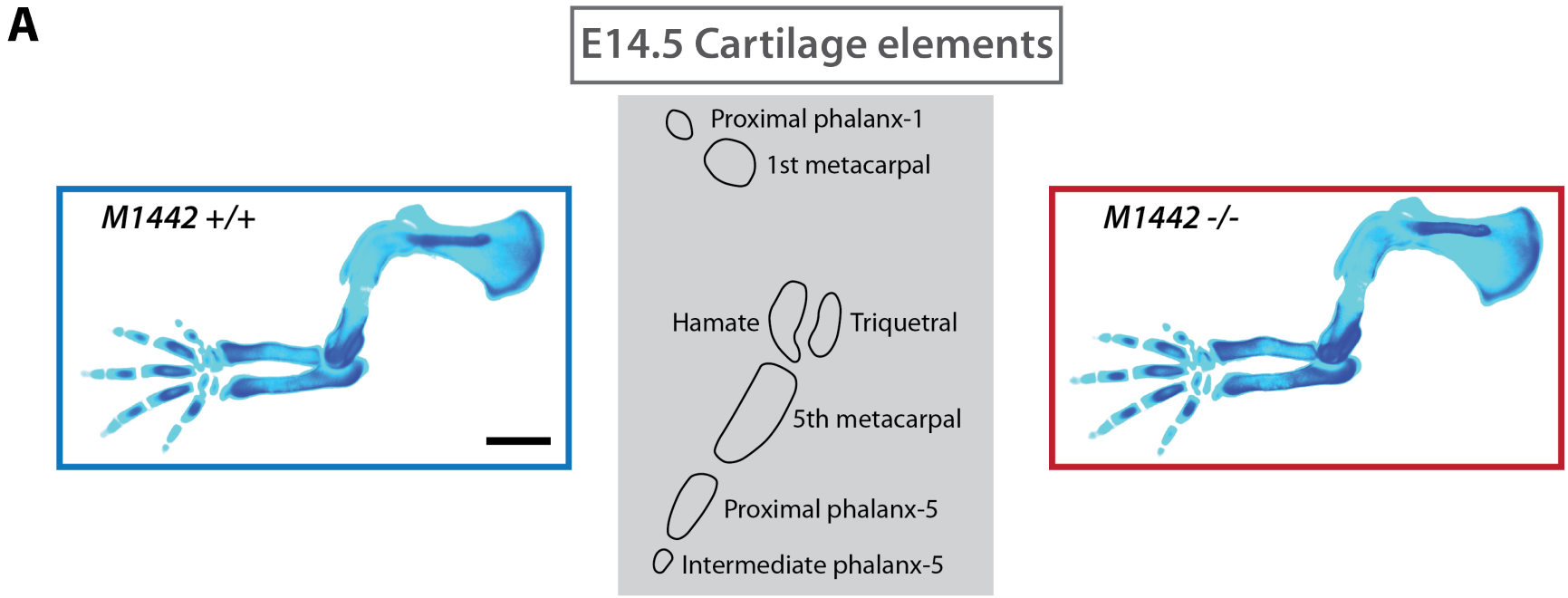
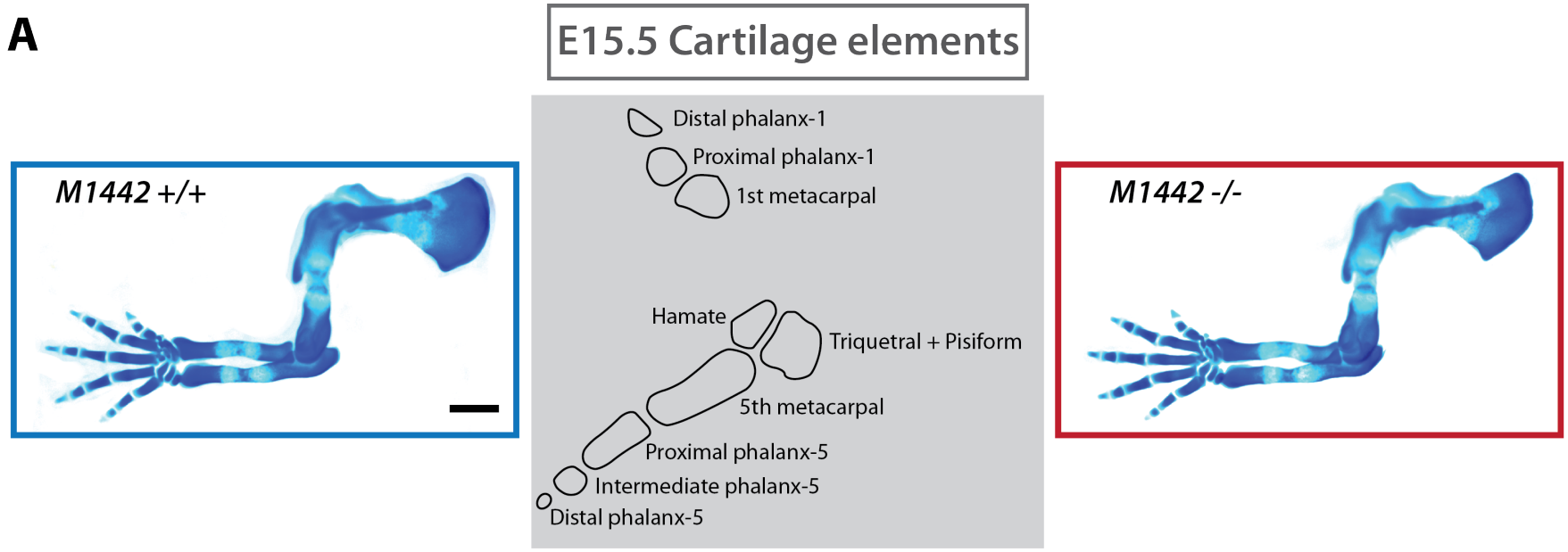

A
E14.5 Cartilage elements
Proximal phalanx-1
1st metacarpal
M1442 +/+
M1442 -/-
Hamate
Triquetral
5th metacarpal
Proximal phalanx-5
Intermediate phalanx-5
B
Mutant
=
<
to/than
control
1
1
1
b
1
a
2
a
2
b
3
a
3
b
1
a
a
b

Figure 36. Forelimb comparisons between *M1442* null and control embryos at E15.5.
(A) Control tissue outlined in blue; mutant tissue outlined in red. Grey box shows cartilage elements used in comparison. (B) Each litter is represented by a different colored box: orange, green or purple. Control limbs (blue) from each litter indicated by Arabic numerals. Mutant limbs (red) from each litter indicated by lowercase letters. If control limb is shorter (and/or more delayed) or larger (and/or more advanced) than mutant limb then it will be indicated by a dotted outline. Scale bars = 0.5 mm.


A
E15.5 Cartilage elements
M1442 +/+
Distal phalanx-1
Proximal phalanx-1
1st metacarpal
Hamate
Triquetral + Pisiform
5th metacarpal
Proximal phalanx-5
Intermediate phalanx-5
Distal phalanx-5
M1442 -/-
B
Mutant
>
=
to/than
control
1
a
1
a
1
b
1
a
a
b
2
1

equal and two were scored as mutant-larger-than-control. The majority of all pairwise comparisons over the three time points were scored as equal (Figs. 34-36).

Assessing phenotypes in adult mutant mice

My research focused on the impact of enhancer deletions on embryonic limb development. However, I did pursue one experiment using adult mice that yielded three important observations about the limbs of adult mutant mice. In addition to this experiment, I made a phenotypic observation unrelated to limb development early in my analysis of *M280*^{tm1} -/- mice that was later confirmed in a statistically rigorous fashion. I will first discuss the experiment conducted on adult limbs and the three resultant observations then proceed to describe the *M280* null-specific phenotype.

Forelimbs from three 6-month-old control, *M280*^{tm1} -/-, and *M1442*^{tm1} -/- female mice were subjected to a micro CT scan and bone mass density (BMD) values and 3-dimensional (3D) surface renderings from each limb were collected. The first important observation from this experiment resulted from my dissecting away the forelimbs from each mouse (I also dissected off hindlimbs as well, but I did not submit them for micro CT scans); I was able to carefully analyze each autopod element as I did in the embryos described above. Additionally, I was able to sketch some of the soft tissue, such as the tendons and muscles. As was the case with my embryonic observations and comparative sketches I did not visually detect any phenotypes in either the hard or soft tissues in the adult mutant forelimbs (and hindlimbs). Secondly, there was no statistical difference in BMD between either control and *M280* null mice or control and *M1442* null mice (Students t-Test: Control and *M280* null, P = 0.68; Control and *M1442* null, P = 0.38; Appendix 1). Thirdly, the 3D renderings produced during the micro CT scans showed that adult *M280*^{tm1} -/- and *M1442*^{tm1} -/- forelimbs were indistinguishable from

control limbs in size and shape (Fig. 37). These observations indicate that limb structure in adult mice is unaffected by the absence of the *M280* and *M1442* enhancers.

Now to the observation I made early on in my analysis of $M280^{tm1}$ mice. To initially generate $M280^{tm1}$ -/- mice I needed to set up heterozygous $M280^{tm1}$ intercrosses. Obviously, pups obtained from these initial crosses would be a mix of $M280^{tm1}$ heterozygotes or homozygotes and *M280* wild types. At weaning age, 3 weeks post-birth, I began noticing that $M280^{tm1}$ -/- males and females were often markedly smaller in overall body size than their heterozygous or wild-type littermates. Though the difference was somewhat subtle it was also clearly obvious that a high percentage of $M280^{tm1}$ null animals were smaller (size difference is discernable in Fig. 13). To rigorously test the hypothesis that $M280^{tm1}$ null mice were smaller, at least in terms of body weight, than either $M280^{tm1}$ heterozygotes or control littermates I recorded the weight of males and females at time of weaning and at six weeks of age, when the animals undergo sexual maturation. All weights were taken at noon so as to reduce weight variation that may result from recordings taken at different times of the day – laboratory mice eat and are active at different times during the day/light cycle. Initially I had planned to use the Student's t-Test to compare body weights of mice at the selected time points from litters of the same size. But I realized very quickly that it was difficult to consistently get litters of the exact size. Not impossible, of course, just difficult in a short time period. I decided, therefore, to submit all of my weight data, regardless of litter size, to a group of dedicated statisticians that could analyze for weight differences and remove confounding differences in weight that may arise due to litter size/milk accessibility. Needless to say, the statistical analysis was a bit more sophisticated than a standard Student's t-Test.

The results of the statistical analysis confirmed my initial observations (Fig. 38A; see Appendix 2 for full report): Female and male $M280^{tm1}$ -/- mice weighed less than control mice at 3 weeks and 6 weeks. Moreover, the weight differences between female and male $M280^{tm1}$ -/-

Figure 37. Micro CT surface renderings of five-week-old forelimb autopods from mutant and control mice. Images not to scale.

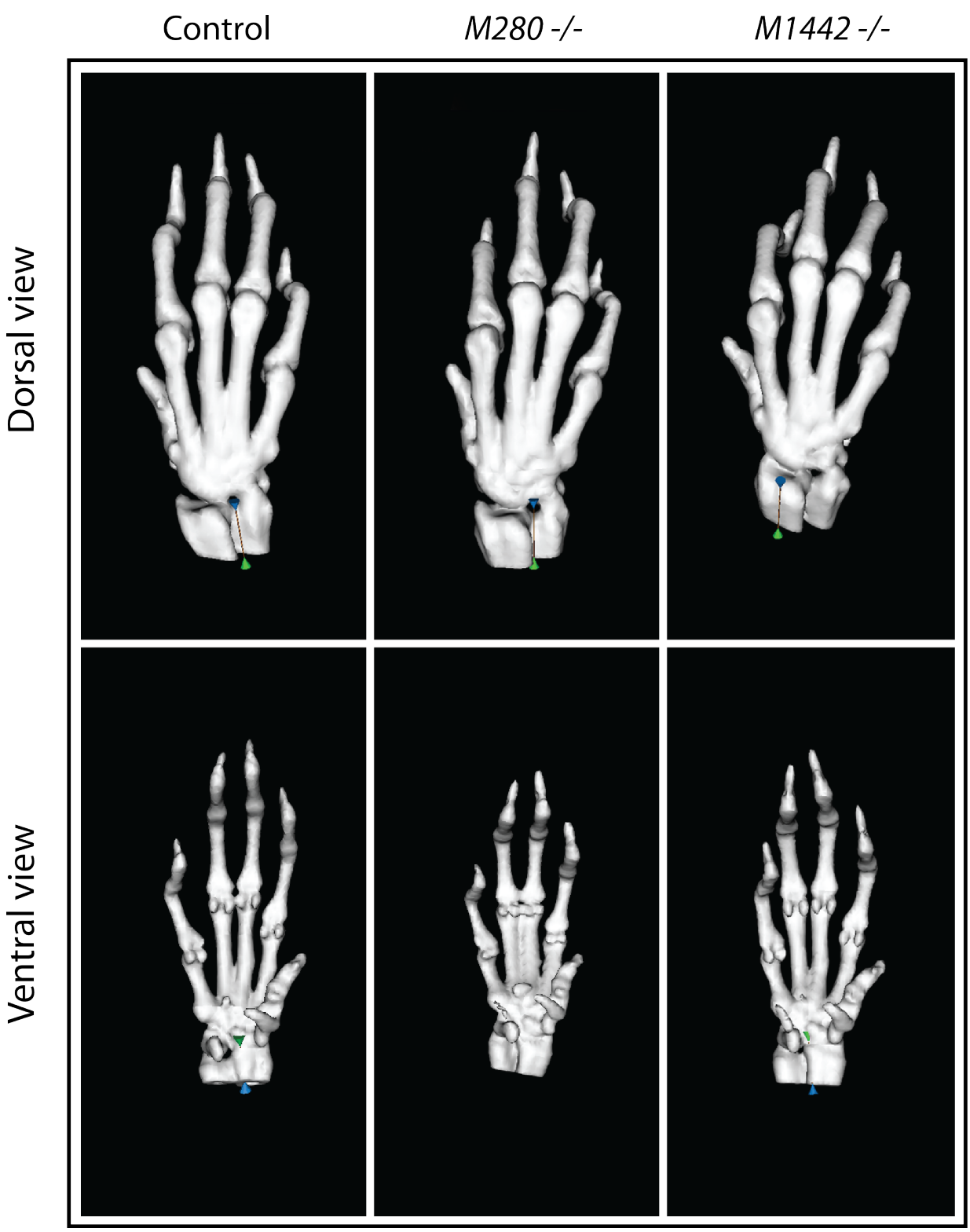
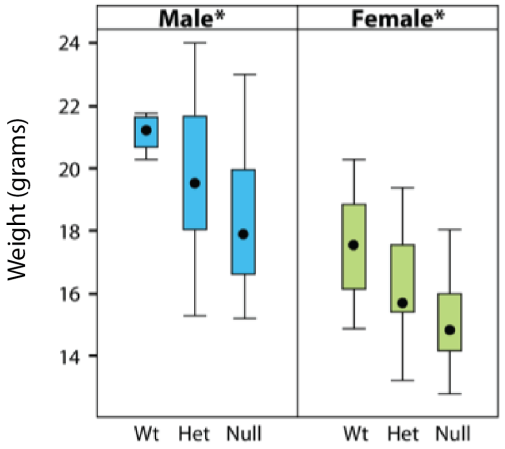

Control
M280 -/-
M1442 -/-
Dorsal view
Ventral view

Figure 38. Adult and embryonic *M280* null animals are smaller than controls. (A) Box plot showing distribution of adult body weight in control and heterozygous and homozygous *M280* mutant mice. (B-F) All other graphs denote either average weights or lengths as specified. Error bars indicate standard deviation. Asterisks indicate significant differences determined via Student's t-Test. See text for details.

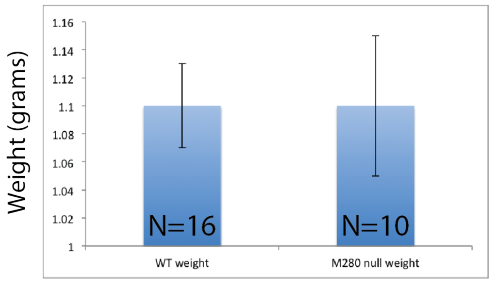
A

Six-week-old Body Weight



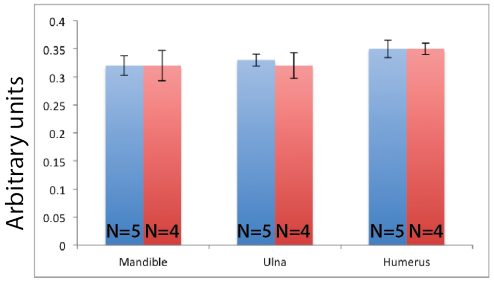
B

E18.5 Body Weight



C

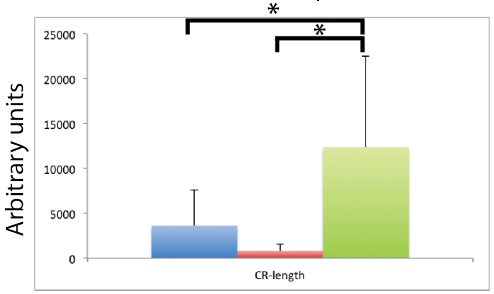
E15.5 Mandible, Ulna, Humerus Length:
Normalized to CR-length



Control: *M280* +/+
M280 -/-

D

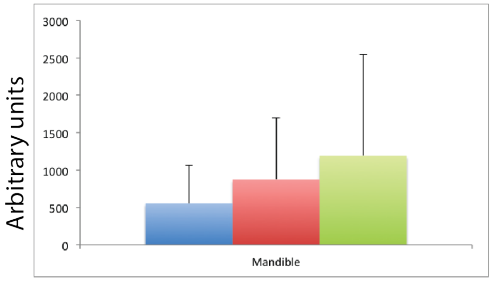
E15.5 Non-normalized CR-length:
Inter-litter Comparison



M280: Null X Null matrix (N=6)
Control X Control matrix (N=10)
Null X Control matrix (N=20)

E

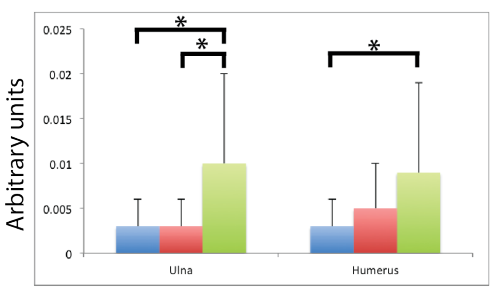
E15.5 Non-normalized Mandible Length:
Inter-litter Comparison



M280: Null X Null matrix (N=6)
Control X Control matrix (N=10)
Null X Control matrix (N=20)

F

E15.5 Non-normalized Ulna, Humerus Length:
Inter-litter Comparison



M280: Null X Null matrix (N=6)
Control X Control matrix (N=10)
Null X Control matrix (N=20)

and *M280*^{tm1} -/+ mice were also either significant or marginally significant, with null mice weighing less that heterozygous mice. Likewise, the weight differences between female and male *M280*^{tm1} -/+ and control littermates were also significant or marginally significant, with heterozygous mice weighing less than controls. In graphical form these results take on a classic null-heterozygous-wild-type geometry, with *M280* null animals showing the most extreme phenotype (lowest body weight relative to control animals), and *M280* heterozygotes manifesting a moderate phenotype, somewhere in between null and control littermates. To be precise, at 3 weeks female and male *M280*^{tm1} -/- mice weigh, on average, about 1 gram (females) to 1.5 grams (males) less than control littermates (control females weigh approximately 10 grams; control males about 10.5 grams). At six weeks male and female *M280*^{tm1} -/- mice weigh, on average, approximately 2.5 grams less than control littermates (control females weigh approximately 18 grams; control males about 21.5 grams; exact numbers available in Appendix 2). Even qualitatively this is a dramatic phenotype. Another way of saying it is that at six weeks, for example, *M280* null mice manifest approximately 10% less body weight than control mice. Interestingly, this phenotype appears to be completely unrelated to limb development or limb morphology as indicated by my careful analysis of embryonic limb development and my micro CT-based analysis of adult limb morphology. Moreover, I collected the body weights of E18.5 mutant and control embryos, all from same-sized litters, and determined that there is no difference in body weight at this late stage of development (N = 16 control embryos; N = 10 *M280* null embryos; Student's t-Test, P = 0.87; Fig. 38B). Therefore, I concluded changes in body weight occur after birth, before weaning, and are amplified by six weeks of age in mutant *M280* mice.

Upon confirming the post-birth weight differences in *M280* null mice I recalled that the transgenic reporter lines *M280*L1 and L2 showed consistent *lacZ* expression in the first pharyngeal arch at E10.5-E11.5; this embryonic structure gives rise to the maxilla and mandible.

I hypothesized, therefore, that perhaps mandible development was perturbed during development and after birth this prohibited *M280* null pups from obtaining nutriment. This could explain the overall decrease in body size in mutant mice. To test this prediction I measured mandible length in mutant and control E15.5 embryos from three litters. I noted no changes in mandible morphology in mutant embryos and when mandible length was normalized to embryo-crown-rump length I observed no differences in mandible length between *M280* and control embryos (Fig. 38C). Therefore, aberrant mandible development does not seem to be causal in producing the body weight phenotype.

While doing the mandible/CR-length measurements I did notice, however, that mutant embryo CR-lengths were always shorter than the control CR-lengths in a given litter. This seemed to dovetail with the observation that overall body weight was decreased in adult mice – was body size in *M280* mutants already stunted in E15.5 embryos? This supposition was in opposition to the fact that I had observed no body weight differences in E18.5 embryos. But this discrepancy could arise for at least two reasons: 1) perhaps E18.5 mutant and control embryos do, on average, weigh the same or at least if there are weight differences they are so minute that they are undetectable; or 2) perhaps the method of preparing the embryos introduces confounding variables that swamp the weight signal. Before weighing the E18.5 embryos I removed the embryonic fluid on their bodies using a Millipore vacuum manifold. However, it is entirely possible that this method did not uniformly dry the embryos; if so, any small weight difference would go undetected.

I was curious to know if the CR-length differences I noted in E15.5 embryos could be quantified. This didn't seem possible at first because the numbers of each genotype, *M280* null and control, in each litter were small and I didn't want to compare CR-lengths across litters due to *in utero* differences that might influence differences between litters (e.g. size of litter; nutriment uptake of mother). But then I devised a way to compare across litters and look for

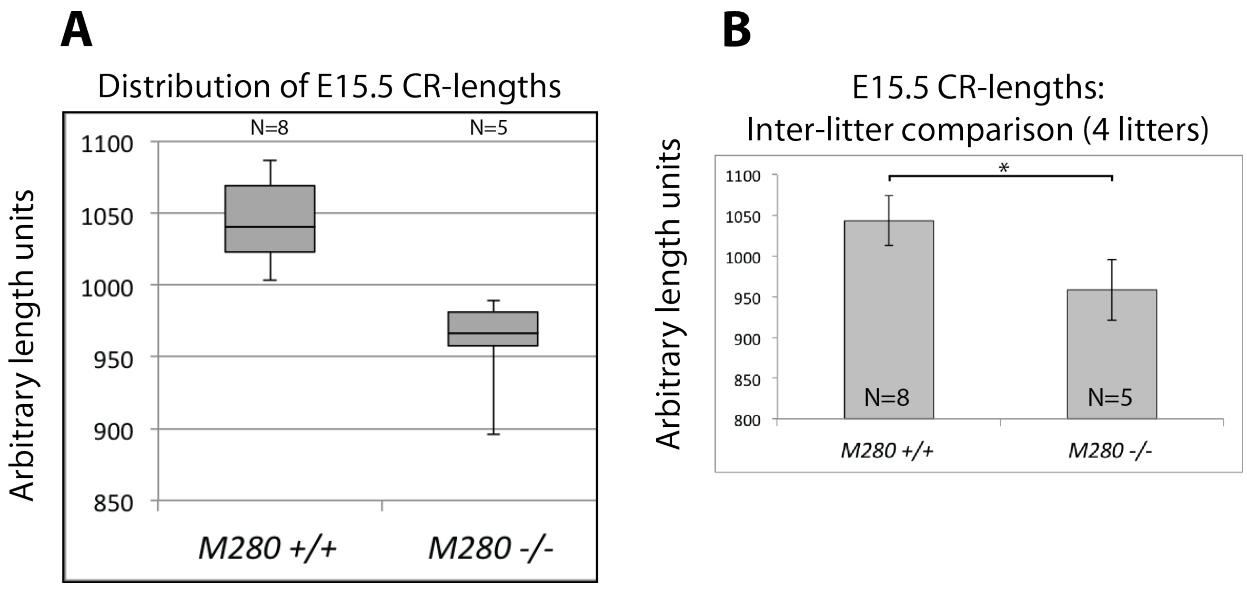
statistical differences in mutant and control CR-length (see examples of matrices in Appendix 1). I created three matrices. In one matrix I calculated CR-length differences between all *M280* null embryos (N=6; NXN matrix), in another differences between all control embryos (N=10; WXW matrix), and in a third matrix the differences between all null and control embryos (N=20; NXW matrix). All difference values were squared to remove negative integers. These matrices allowed me to account for CR-length variation across all litters (*in utero* confounding variables, for example) and still assess whether differences existed between mutant and control animals, and to obtain enough independent observations to perform statistics. I predicted that *M280* null CR-length was less that control CR length in a statistically significant manner. To test this I performed a Student's t-Test between all matrices (assuming unequal variance; F-test indicated unequal variance between all matrices: NXN v. WXW, $P=9.2 \times 10^{-5}$; WXW v. NXW, $P=5.8 \times 10^{-9}$; NXN v. NXW, P=0.04). If indeed *M280* null CR-lengths cluster together after accounting for variation in body size from litter to litter, and if the same is true for control CR-lengths, then the average NXN value should not be statistically different that the average WXW value. This is what I see after carrying out a Student's t-Test between NXN and WXW (P=0.14; Fig. 38D). On the other hand, if there are consistent differences in CR-length between E15.5 mutant and control embryos then the average NXW value (that is CR-length differences between all control and mutant embryos across three litters) after accounting for across-litter variation should be significantly different than the average NXN or WXW values. This was also the case; differences in CR-length between mutant and control embryos are larger than CR-length differences between either all null embryos or all control embryos (Student's t-Test: WXW v. NXW, $P=6.2 \times 10^{-5}$; NXN v. NXW, P=0.0046; Fig. 38D). From these results I concluded that *M280* animals already exhibited delayed body growth at E15.5. This means that by E18.5 either mutant animals have compensated for this delayed growth and then succumbed to it again by three weeks of age, or the delayed growth phenotype is present at E18.5 but I did

not detect it. These results also help explain why the majority of mutant forelimbs in E15.5 cartilage comparisons were shorter than controls – the whole E15.5 embryos is smaller (Fig. 33).

Additional matrix-based analysis was done to compare other cartilage elements in E15.5 *M280* and control embryos from three litters. Not all of these comparisons indicated that E15.5 mutant embryos were smaller than controls, but numerous comparisons were significant. Raw lengths for the ulna and humerus were shorter in mutant embryos than in controls when accounting for variation across litters (but this was not true for the mandible; Fig. 38E, F). However, when ulna and humerus lengths were normalized against CR-lengths there was no difference between mutants and controls (Fig. 38C). Again this demonstrates that the whole mutant animals, not just their individual limb elements, are smaller than control animals at E15.5.

Lastly, the statistical analysis performed on adult body weights discussed earlier found no significant correlation between litter size and weight differences between genotypes (Appendix 2). Thus, I assumed that litter size does not confound comparisons of CR-length across litters – just as litter size did not confound body weight in adults. Therefore, I directly asked whether mutant *M280* CR-lengths are significantly shorter than those of controls using just the raw CR-length values (not using the matrices described above). Embryos for this direct comparison were drawn from four litters generated by $M280^{tm1}$ -/+ intercrosses. I found that CR-lengths in E15.5 *M280* mutants are significantly shorter than control CR-lengths (Fig. 39A, B; Student's t-Test, $p<0.001$). Also concordant with these findings is that raw, non-normalized mandibular and ulnar lengths are shorter in *M280* mutants than in controls; and the lengths of mutant humeri are also generally shorter than controls (Appendix 1). Therefore, we conclude from these observations that at E15.5 whole *M280* embryos are smaller than control embryos, but that the proportional length of individual bone elements to overall body size, such as the

Figure 39. E15.5 *M280* null animals are smaller than controls. (A) Box plot showing distribution of E15.5 CR-lengths in *M280* null and control embryos. (B) Graph showing average E15.5 CR-length in *M280* null and control embryos; error bars indicate standard deviation. Asterisk indicates significant difference determined via Student's t-Test. See text for details.


A
Distribution of E15.5 CR-lengths
N=8
N=5
1100
1050
1000
950
900
850
Arbitrary length units
M280 +/+
M280 -/-
B
E15.5 CR-lengths:
Inter-litter comparison (4 litters)
*
1100
1050
1000
950
900
850
800
Arbitrary length units
N=8
N=5
M280 +/+
M280 -/-

mandible or ulna, are maintained in the mutants. Thus the weight differences observed between mutant and control postnatal mice are presaged by embryonic size differences by at least E15.5.

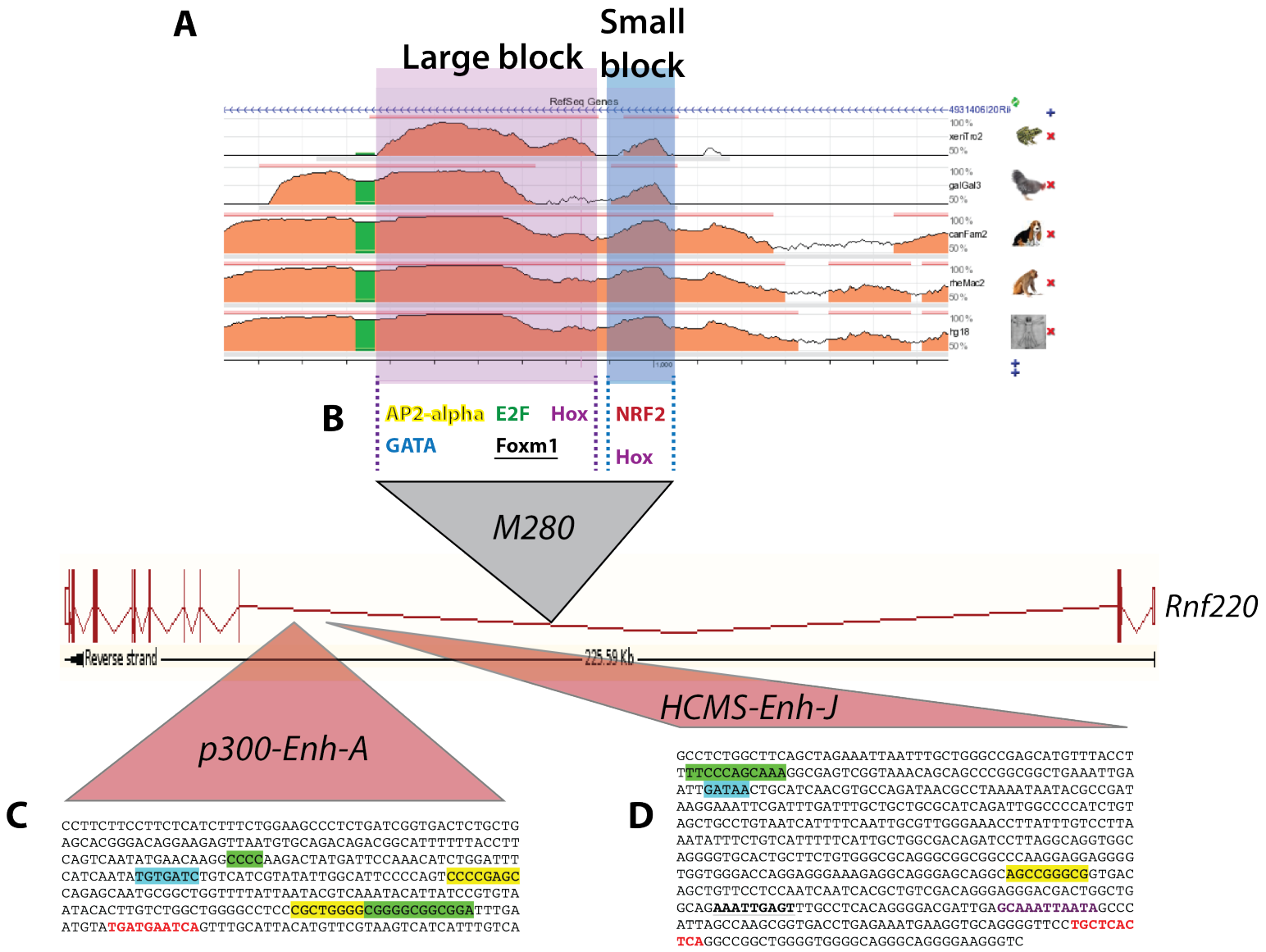
These results underscore several important concepts in enhancer biology. Firstly, just because an enhancer regulates gene expression in a particular embryonic tissue, such as the limb, does not mean that its main function is to direct gene expression in that tissue during embryonic development. Such an enhancer may actually contribute to ordinary development before the time at which it is assayed (in a transgenic reporter study, for example) and/or contribute to tissue maintenance post-birth regardless of its regulatory potential during development. A corollary to this concept is the notion that enhancers with tissue-specific regulatory capacity during development may in fact be required in adult tissues that are completely different from the embryonic regions used to characterize the enhancer in the first place. My observations described above are at the very least suggestive of this concept. Enhancer *M280* was characterized mainly as a limb-specific enhancer in transgenic reporter gene assays, but the result of its deletion from the mouse genome – reduced embryonic size and post-birth body weight – appears to be unrelated to its ability to drive gene expression in embryonic limbs. Presumably, the cells in some other tissue/organ in maturing pups that is not the limb (again, presumably, because I have not looked for *lacZ* expression in adult limbs) also expresses the proper milieu of *M280*-enhancer-binding transcription factors to engage *M280* with its target gene(s); and presumably this *M280*-guided transcription is required for proper mid-gestation and post-natal development or maintenance of body weight.

In silico approach to identifying enhancers that may compensate for the deleted M280 enhancer

Numerous studies indicate that multiple enhancers, often positioned in relatively close genomic proximity, direct gene expression to the same region. Such enhancers are often called shadow enhancers because they mimic, or shadow, the action of one another. Even in experiments where no such shadow enhancers have been experimentally validated, their presence had been postulated because the disruption of a particular enhancer does not lead to an expected phenotype or the phenotype is of a subtle nature. One potential mechanistic explanation for this phenomenon is that each of these shadow enhancers possesses a similar combination, orientation and spacing of TF binding sites. Thus, in cells expressing the right mix of TFs all of the shadow enhancers should be able to contribute to the activation of their target gene(s). In theory then, it should be possible to identify shadow enhancers based on their composition of TF binding sites relative to an enhancer of interest. Following such logic I predicted that I could use *M280*'s combination of TF binding sites to locate *M280*-shadow enhancers within a restricted genomic window that encompasses *M280*.

I submitted the entire *M280* sequence to the ECR Browser, freely available software that aligns vertebrate genomic sequences and scores their level of conservation. After submission I noted that *M280* could be divided into two conservation blocks, one large 520 bp block and one small 148 bp block. Core regions within the large and small blocks were conserved from mice to frogs. Transcription factor binding sites in these two blocks were obtained using rVista (rvista.dcode.org). Eleven of the 184 returned TF binding sites corresponded to TFs that had an association with limb development in previous publications. The *M280*-embedded TF binding sites include those for AP2α, E2F, Clox (Cux1), ETS, Hox TFs, GATA and Foxm1. These TF binding site motifs were used to identify other potential shadow enhancers ± 1 Mb from *M280*. Two putative shadow enhancers, drawn from a list of highly conserved and p300-associated putative enhancers, shared the same cluster of TF binding sites with *M280* (Fig. 40). Interestingly, both these putative shadow enhancers are located within the same large second

Figure 40. Putative shadow enhancers reside in the same *Rnf220* intron as *M280*.
(A) Two sequence blocks within *M280* evolutionarily conserved from mammals to anurans (*Xenopus*). Image generated on the ECR Browser (ECRbrowser.dcode.org) (Ovcharenko et al., 2004): green = simple repeats; salmon = intronic regions. (B) Transcription factor (TF) binding sites within each conserved block identified in rVISTA. Color-coded TF binding sites in (B) correspond to colors in (C). (C) Two putative shadow enhancers obtained from previous screens for enhancers based on bioinformatics and p300-ChIP-seq datasets (Visel et al., 2008; Visel et al., 2009a).


A
Large block
Small block
RefSeq Genes
xenTro2
galGal3
canFam2
rheMac2
rg18
B
AP2-alpha
E2F
Hox
NRF2
GATA
Foxm1
Hox
M280
Rnf220
Reverse strand
225.59 Kb
HCMS-Enh-J
p300-Enh-A
C
CCTTCTTCCTTCTCATCTTTCTGGAAGCCCTCTGATCGGTGACTCTGCTG
AGCACGGGACAGGAAGAGTTAATGTGCAGACAGACGGCATTTTTTACCTT
CAGTCAATATGAACAAGGCCCCAAGACTATGATTCCAAACATCTGGATTT
CATCAATATGTGATCTGTCATCGTATATTGGCATTCCCCAGTCCCCGAGC
CAGAGCAATGCGGCTGGTTTTATTAATACGTCAAATACATTATCCGTGTA
ATACACTTGTCTGGCTGGGGCCTCCCGCTGGGGCGGGGCGGCGGATTTGA
ATGTATGATGAATCAGTTTGCATTACATGTTCGTAAGTCATCATTTGTCA
D
GCCTCTGGCTTCAGCTAGAAATTAATTTGCTGGGCCGAGCATGTTTACCT
TTTCCCAGCAAAGGCGAGTCGGTAAACAGCAGCCCGGCGGCTGAAATTGA
ATTGATAACTGCATCAACGTGCCAGATAACGCCTAAAATAATACGCCGAT
AAGGAAATTCGATTTGATTTGCTGCTGCGCATCAGATTGGCCCCATCTGT
AGCTGCCTGTAATCATTTTCAATTGCGTTGGGAAACCTTATTTGTCCTTA
AATATTTCTGTCATTTTTCATTGCTGGCGACAGATCCTTAGGCAGGTGGC
AGGGGTGCACTGCTTCTGTGGGCGCAGGGCGGCGGCCCAAGGAGGAGGGG
TGGTGGGACCAGGAGGGAAAGAGGCAGGGAGCAGGCAGCCGGGCGGTGAC
AGCTGTTCCTCCAATCAATCACGCTGTCGACAGGGAGGGACGACTGGCTG
GCAGAAATTGAGTTTGCCTCACAGGGGACGATTGAGCAAATTAATAGCCC
ATTAGCCAAGCGGTGACCTGAGAAATGAAGGTGCAGGGGTTCCTGCTCAC
TCAGGCCGGCTGGGGTGGGGCAGGGCAGGGGAAGGGTC

Rnf220 intron as *M280* (*M280-pSh1* and *M280-pSh2*, respectively; pSH = putative shadow enhancer). The putative enhancers are only separated by 2.3 kb, and are jointly approximately 74.5 kb from *M280*. If common TF binding site signatures can predict common enhancer function, then it is possible that *M280-pSH1/2* shares at least an overlapping regulatory roll with *M280*; thus the action of this putative shadow enhancer pair may dampen any limb phenotype would otherwise result from *M280's* deletion.

DISCUSSION

Mouse enhancer *M280* influences embryonic and adult body size: the first phenotype reported for a deleted ultraconserved element

Although the decision to delete *M280* from the genome was based on its ability to drive reporter gene expression in embryonic limb buds the results reported above demonstrate that it is required for the development of embryonic body size and adult body weight. Thus, though *M280* certainly directs gene expression in particular tissues, such as the limb buds and first pharyngeal arch, its effects on organismal development are global in nature. The details of the molecular mechanism behind the size reduction that accompanies the targeted deletion of *M280* are not clear at this point. Nor is it clear whether *M280*'s role in body size production and/or maintenance is a result of its gene regulatory capacities throughout the entire organism or its role in one tissue, such as the liver or the placenta. The analysis of *M280*'s regulatory functions encompassed the whole external embryo, as well as a special focus on both the external and internal features of the limb. In future work it would be interesting to thoroughly examine the internal organs and extra-embryonic tissues in which *M280* can drive reporter gene expression; such an analysis will likely provide clues as to when, where and how *M280* is regulating gene expression essential to body size development. It is worth noting that numerous previous studies have linked genes involved in placental development with reduced adult body size in mutant mice (Li and Behringer, 1998; Rodriguez et al., 2004; Murthi et al., 2012); thus it will be especially interesting to assess *M280*'s ability to regulate gene transcription in the placenta (particularly between E14.5-15.5 when the beginnings of decreased body size in mutant *M280* mice was noted).

Importantly, along with its affiliation with the general transcriptional activator, p300, *M280* also spans an ultraconserved element, which is defined as a sequence of ≥ 200 bps of perfect conservation between human, rat and mouse (Bejerano et al., 2004; Visel et al., 2008). To date, only four other ultraconserved elements had been deleted from the mouse genome (Ahituv et al., 2007). These four other ultraconserved elements were selected because of their proximity to genes that have been shown to be required for particular developmental processes in knockout mice experiments. However, none of the phenotypes resulting from the targeted deletion of the genes is recapitulated in the ultraconserved element-knockout mice (Ahituv et al., 2007). Therefore, the significant decrease in body size and weight noted in *M280* null mice is the first phenotype reported for a deleted ultraconserved element; these results make the case that ultraconserved elements encode regulatory abilities necessary for the normal development of vertebrates.

Additionally, these results suggest that the phenotypic consequences of manipulating an ultraconserved element, or perhaps any enhancer, may be unrelated to their capacities to drive expression in a particular tissue. This general observation should give pause to future researchers engaged in endogenous enhancer-deletion studies that don't immediately score a phenotype in an expected tissue. An enhancer's regulatory capacities in transgenic contexts – wherein the enhancer and an associated reporter gene are inserted randomly into the genome – may not recapitulate its role in an endogenous context. Thus conclusions drawn from transgenic reporter gene assays can be completely informative, partially informative, partially misleading, and possibly, entirely misleading; hence the importance for using gene targeting or similar techniques to delete or otherwise manipulate enhancers *in situ*.

***M1442* and *M280* not required for embryonic limb development**

The decision to assess the role of *M280* and *M1442* during mouse limb development stems from two initial observations surrounding their identification and characterization: 1) in mouse limb buds both these sequences were associated with p300, which is known to be associated with active enhancers and contributes to the acetylation of H3K27, a chromatin modification that is a signature of active transcription; and 2) their orthologous human sequences directed reporter gene expression in the limbs of E11.5 transgenic mice. I demonstrated that *M280* and *M1442,* like their human sequence counterparts, direct gene expression in mouse limb buds. Importantly, I showed that transcription regulated by these enhancers begins before E11.5 (both *M280* and *M1442*) and extends beyond this time as well (*M1442*). One important conclusion from this research, therefore, is that enhancers identified at specific time point in a specific tissue may well have regulatory capacities that extend beyond the criteria that led to their discovery.

As discussed above, *M280* is associated with p300 and harbors a 256 bp ultraconserved element. Moreover, its sequence identity along its entire length when compared to the homologous human sequence is 94.9% (UCSC Genome Browser, mm9 and hs19). Similarly, *M1442* is associated with p300 and its sequence identity along a core 1,440 bp region that aligns to the homologous human sequence is 91.9% (UCSC Genome Browser, mm9 and hs19; more concerning the non-homologous sequences below). Both *M280* and *M1442* (and their human counterparts) drive *lacZ* expression in the developing limb. Despite these observations neither *M280* nor *M1442* are required for mouse limb development. In the absence of *M280* and *M1442* limbs developed normally, without even transient developmental defects, all limb bone elements were accounted for, and right before birth limb bone lengths were indistinguishable between mutant and control animals. Much of this work focused on assessing the effect that

M280/M1442 absence might have on hard tissue development. However, in histological sections, adult dissections, and in the day-to-day observations of the mutant animals I did not see any indication that limb soft tissue was faulty. The lack of a limb phenotype in the *M280* and *M1442* null mice represent only the second and third examples, respectively, of an undetected phenotype when a limb enhancer is deleted via gene targeting (Fig. 1). The first example, already discussed, showed that *Prx1*-limb-enhancer null mice manifested no tractable limb defects and that *Prx1* transcription was unaltered by the absence of this enhancer (Cretekos et al., 2008). Hence another important general conclusion drawn from this research is that enhancers identified via their association with the chromatin modifying enzyme, p300, appear to be no more likely than enhancers discovered via traditional methods (proximity to genes and sequence conservation) to be required for limb development.

The results of the *M280* knockout experiment are relevant for another reason. Recently, Cotney *et al.* (2013) demonstrated that *hs280*, the homologous region in the human genome to *M280*, had increased H3K27ac relative to orthologous sequences in rhesus macaques and mice (Cotney et al., 2013). Authors of this study suggest that the increased H3K27ac specifically in the human lineage, like those observed at *hs280*, correspond to increased gene expression and thus to transcriptional differences that may account for the evolution of human-specific traits. The deletion of *M280* in the mouse genome indicates, however, that neither sequence conservation nor increased H3K27ac during embryonic development nor ability to drive reporter gene expression can predict whether an enhancer is necessary for the generation of necessary anatomies and organs, such as the limb. I am not concluding that the convergence of multiple tests of importance for a particular enhancer are useless, but I am suggesting that our current understanding of what differentiates enhancers from other genomic elements and from each other is not sufficient to predict their role during embryonic or postnatal development.

M280 -/- and *M1442* -/- mice were not entirely without observable phenotypes. I demonstrated that minor, though statistically significant transcriptional changes were detected within the vicinity of each enhancer. I will discuss the transcriptional changes near *M280* first, and then those observed near *M1442*.

Although *M280* is located within the second intron of *Rnf220* its expression is not affected by the absence of *M280.* This is not surprising, as limb enhancers have been shown to reside both within the genes they regulate and within the introns of neighboring genes. What is interesting is that the transcription levels of *Tmem53* and *Dmap1* are reduced in E11.5 limbs, demonstrating that an enhancer's regulatory effects can span quite large genomic distances, a total of 400 kb in the case of *M280*. Moreover, it is interesting to note that *Eri3*, which lies between *M280* and *Dmap1* does not experience a drop in transcription in the absence of *M280*. This transcriptional observation suggests a 3-dimensional structure to the DNA in this genomic interval. One can imagine that the chromatin in between *M280* and *Dmap1* must undergo a looping event to bring *M280* into the vicinity of *Dmap1*'s promoter while skipping *Eri3*'s promoter, just like the ends of a jump-rope would move closer together if the middle of the rope were lifted vertically. Of course it is not possible to say whether *M280*'s main regulatory role is to promote *Tmem53* and *Dmap1* transcription because no distinctive limb-specific expression pattern of these genes via whole mount *in situ* hybridization. It is possible that *M280*'s regulation of these genes is secondary to a more central role in regulating a gene outside the assayed genomic window, perhaps one that does possess a central limb-specific expression pattern.

Near the *M1442* locus, *Nme7*, which is approximately 500 kb 5' of *M1442,* experienced a 40% decrease in transcription in $M1442^{tm1}$ -/- embryonic limbs. WMISH for *Nme7* did not indicate a strongly demarcated limb expression pattern for this gene, but its general expression does appear to be slightly decreased in *M1442* limb buds. Again, it is not possible to determine

that this regulatory effect is *M1442*'s primary role – it is entirely possible that a gene(s) outside the outside window is regulated by *M1442* in a manner consistent with our transgenic mice. Of interest at the *M1442* locus is the observation that three genes that span a 600 kb window show a slight increase in transcription levels, including *Kifap3*, which hosts the *M1442* enhancer (the other genes are *Gorab and Scyl3*). This is a somewhat unexpected result because transgenic mouse experiments showed that *M1442* should activate transcription, not repress it. One interpretation of these results is that it is not the absence of a repressor element within *M1442 per se* that is causing the minor increase in transcription, but the topographical shift of the promoter regions of the affected genes (Spitz et al., 2005). Perhaps the removal of the approximately 3 kb *M1442* sequence moved the promoter regions of these genes nearer to another enhancer element. Alternatively, it is possible that *M1442* harbors repressor and enhancer sequences and in its endogenous context the repressor elements override the enhancer elements for at least *Kifap3*, *Scyl3* and *Gorab* (*M1442* also possesses flanking sequences that are not homologous to *hs1442*; these non-homologous sequences potentially encode repressor elements). Yet one more possibility is that additional enhancer sequences must act in concert with *M1442* to properly regulate transcription in the genomic window and in the absence of *M1442* this balanced regulatory interaction is thrown off and a slight up-regulation of nearby genes results.

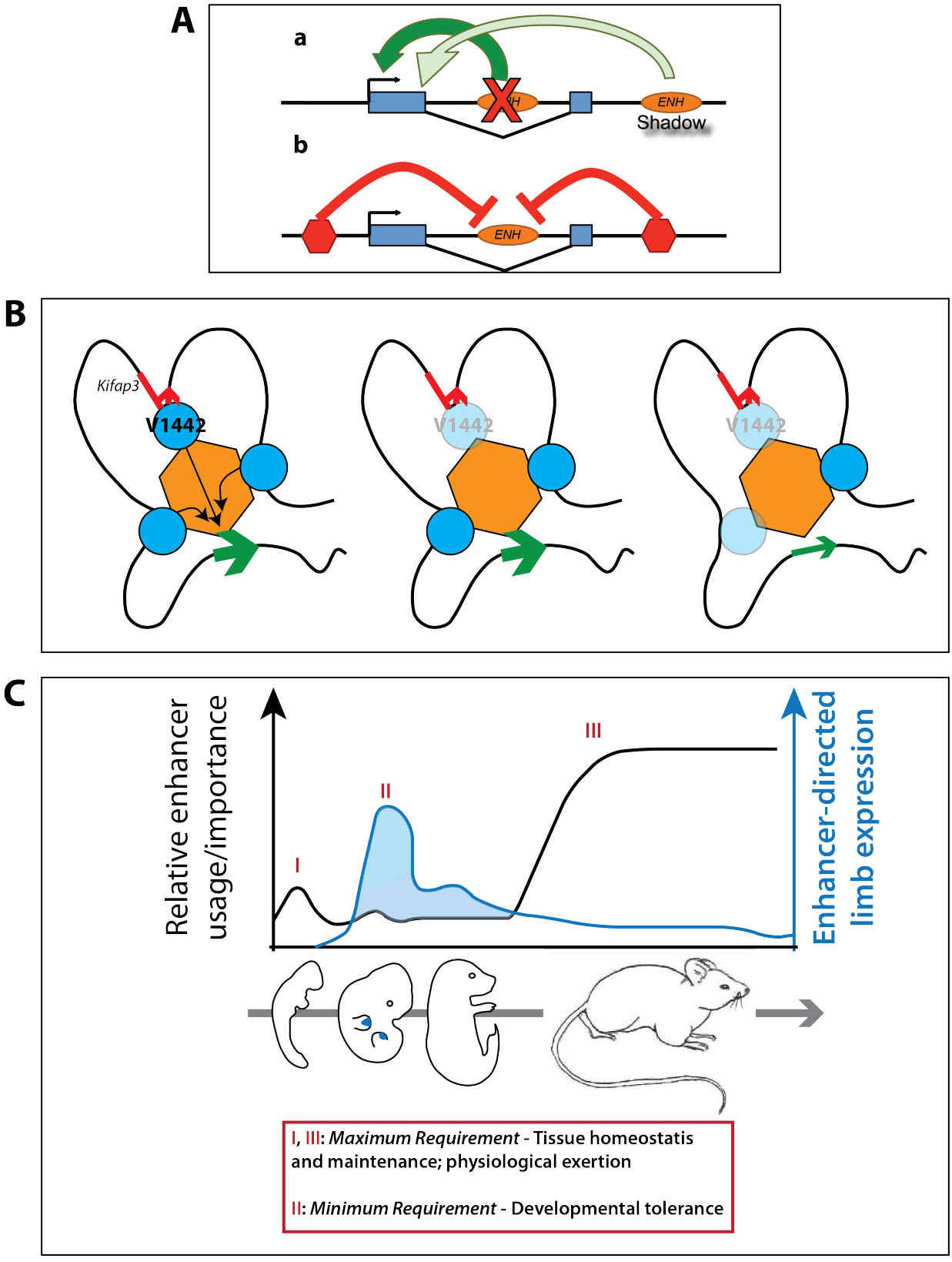
Potential explanations for the lack of limb phenotypes in *M1442* and *M280* mutant mice

Gene transcriptional alterations near the *M280* and *M1442* enhancer deletions ranged from 30% increases to 45% decreases. Similar transcriptional alterations associated with other limb enhancer loci have been correlated with observable limb phenotypes. In particular, a malformed digit V accompanied a deletion of a particular enhancer region, which decreased the

limb-specific transcription levels of *Hoxd13* by 60 % (Montavon et al., 2011). Therefore it is inappropriate to suggest that the limbs of *M280* and *M1442* mutant animals are indistinguishable from controls because the qPCR results indicate that transcriptional changes were modest. Rather, the difficulty of scoring limb phenotypes in *M1442* and *M280* mutant mice is likely due to one or more complex molecular and environmental conditions explored below.

A common explanation for the lack of phenotypes, or the presence of only minor phenotypes, in enhancer knockout experiments is that additional enhancers compensate for the loss of those that were deleted (Fig. 41A.a). This explanation is invoked in conceptual treatments of *cis*-regulatory elements as well as in empirical studies that record the discrepancy between the regional expression pattern driven by a particular enhancer and the absence or subtly of phenotypes in the same region in enhancer-knockout animals. Certainly, distinct enhancers, often referred to as shadow enhancers, capable of driving similar and overlapping expression patterns have been well documented in multiple species, from fruit flies to mice (Levine, 2010; Montavon et al., 2011; Marinic et al., 2013). Also, the evidence for networks of enhancers that contribute to the expression of a single gene or gene groups is incontrovertible. Perhaps the best-documented enhancer networks in limb development are those composed of numerous interacting enhancer regions 5' and 3' of the *Hoxd* cluster. Deletion of swaths of these enhancer regions result in aberrant autopod development in the mouse, but perhaps not as aberrant as initially expected. As just one example, Montavon, T. *et al.* (2011) deleted a 540 kb region 5' of the *Hoxd* cluster that contains four transgenic-reporter-mouse-validated limb enhancers and recorded only a slight length reduction of digits II and V (Montavon et al., 2011). Though four well-characterized *Hoxd* cluster enhancers were deleted, additional enhancers in the network, 5' of the deleted region, apparently compensated partially for the missing regulatory elements and softened the phenotypic impact. This example highlights both the

Figure 41. Models of enhancer behavior. (A) Enh = enhancer. (B) Hexagon = basal transcriptional machinery; blue circles = enhancers in same regulatory network. The green arrow represents the network-regulated gene; thickness of arrow indicates fidelity of transcription. *V1442* = Vista*1442* in reference to the human sequence (*hs1442*) on the Vista Enhancer Browser (enhancer.lbl.gov), and could also be *M1442* in this study or any other ortholog of this sequence. Diagram based on figures in Montavon, T. *et al.* (2011) (Montavon et al., 2011). (C) Light blue shading highlights discrepancy between extra-regulatory capacity of an enhancer and its actual role in development or postnatal life.


A
a
ENH
Shadow
b
ENH
B
Kifap3
V1442
V1442
V1442
C
Relative enhancer usage/importance
I
II
III
Enhancer-directed limb expression
I, III: Maximum Requirement - Tissue homeostatis and maintenance; physiological exertion
II: Minimum Requirement - Developmental tolerance

strengths and weaknesses of arguments that invoke redundant enhancers as an explanation for absent or subtle phenotypes in enhancer knockout mice.

The strength of such an argument lies in the interpretation that enhancer network elements may be partially redundant; this would explain the observation that noticeable phenotypes do accompany enhancer deletions, but such phenotypes do not span the entire anatomical domain within which the enhancer is able to operate, as shown through transgenic reporter assays, for example. One interpretation of current empirical evidence also suggests that partial regulatory redundancy may be common (Fig. 41B), while full redundancy likely is not: to date seven enhancers/enhancer regions known to regulate genes involved in limb development have been endogenously deleted, and in five of these six situations limb phenotypes were documented (Table 1) (Zuniga et al., 2004; Sagai et al., 2005; Cretekos et al., 2008; Feng et al., 2008; Menke et al., 2008; Montavon et al., 2011; Marinic et al., 2013). This again suggests full redundancy is rarely, if ever, built into the regulatory structure of genomes.

Therefore, the weakness of enhancer redundancy arguments lies in the interpretation that separate enhancers can completely compensate for each other; in other words, that one enhancer can be totally redundant with another. Though this extreme interpretation of the redundant enhancer model is rare, it can be easily invoked in situations where no detectable phenotypes accompany enhancer deletions – situations exactly like the one demonstrated herein. The reason for mistrusting the presence of fully redundant enhancers is because it is difficult to imagine situations in which fully redundant enhancers would evolve in the first place. If two genetic elements perform the same function in a completely redundant fashion then one or both elements will likely be open for accumulation of mutations, both adaptive and maladaptive – hence the observation of pervasive pseudogenes in genomes.

Moreover, theoretically, rigidly conserved enhancer sequences are the least likely to be maintained along with redundant enhancers over evolutionary time. *M280*, for example, is

highly conserved in vertebrates, even comprising an ultraconserved element (256 bps of perfect conservation between mouse, rat and human). Therefore, if its regulatory action needed to be backed up by the evolution of a fully redundant shadow enhancer then it would be expected that *M280* frequently acquired mutations that would compromise its regulatory ability. But in fact, the opposite is observed; *M280* is highly conserved. Following this reasoning, highly variable enhancers (perhaps identified via pro-transcription chromatin signatures rather than sequence conservation) would be the enhancers most likely to evolve alongside redundant shadow enhancers. Thus a conceptual evolutionary paradox explains one reason why assigning fully redundant enhancers as the explanation for the lack of phenotypes in the limb enhancer knockout experiments of this study is dissatisfying: only enhancers prone to frequent mutation would be expected to evolve alongside shadow enhancers that completely compensate for their regulatory activity, but the very observation of high conservation in enhancers such as *M280* and *M1442* suggests the opposite; it suggests that they are either not subject to frequent mutation for some genomic-architectural reason or that mutations within them are strongly selected against (Halligan et al., 2011). Highly conserved enhancers are unlikely to work in concert with fully redundant enhancers precisely because their very conservation precludes them from requiring the parallel evolution of a fully redundant enhancer.

This discussion about the semantics of enhancer redundancy is valuable because it permits one to move beyond a conclusion that only posits regulatory redundancy as an explanation for the fact that limb development is unaffected in mice lacking enhancers that most certainly possess limb-specific regulatory capacities. Other equally valid explanations may not be as straightforward but they integrate the whole of the data produced from *M280* and *M1442* null mice and they hint at potential experimental paths that could lead to a better understanding of enhancer identification and function.

One possible explanation for the lack of a limb phenotype in *M280* and *M1442* mice is that these sequences can actively promote transcription when removed from their endogenous loci, but *in situ* their limb-regulatory properties are largely suppressed (Fig. 41A.b). Perhaps adjacent to *M280* and *M1442* are additional regulatory elements that override their ability to regulate transcription in the limb domain. One way to test this possibility would be to use gene targeting to knock in a minimal promoter and reporter gene, such as *HSP68* and *lacZ*, respectively, adjacent to either *M280* or *M1442*. If the reporter gene were silent throughout limb development one interpretation would be that at their endogenous loci *M280* and *M1442* are acted on by other regulatory elements in such a way that their implicit limb-regulatory capacity is inhibited. It has been shown that an enhancer's activity is different depending on whether it is assayed at endogenous or ectopic genomic locations. The transposon-based enhancer sensor, SBlac, is distributed throughout the mouse genome. When proximal to enhancer elements *lacZ* is expressed from the SBlac sensor. Ruf *et al.* (2011) demonstrated that when SBlac is positioned near enhancers characterized via their association with p300, the expression pattern of *lacZ* from the transposon is often more restricted than when a p300-enhancer (such as *M280* or *M1442*) drives *lacZ* in an ectopic transgenic context (Ruf et al.).

Another explanation for the observed results could be that numerous enhancers, though promoting transcription throughout embryogenesis, may in fact not be required for development but rather for an organism's adult life. Certainly there are numerous exigencies experienced in a wild adult mouse lifespan that may require genetic input – these may include the physiological responses to the drives for and stress of finding a mate, finding shelter and food, avoiding predation, and defending and patrolling a territory. Physiological exertions required to meet these demands would take a toll on tissues such as muscles and bones. Tissue homeostasis is also an important cellular process that is likely taxed in non-laboratory conditions – certainly genetic switches and pathways that determine when and where to allocate metabolic resources

in response to stress and other environmental cues are an important aspect of an organism's fitness. Concordant with this reasoning then is the possible conclusion that *M280* and *M1442* (and other deleted enhancers that have no observed phenotypic effect such as *UC248, UC329, UC467* and *UC482) (Ahituv et al., 2007)* are developmentally tolerated enhancers. That is, *M280* and *M1442* can promote transcription in embryonic limbs, but that is secondary to their role in the adult mouse (or in the regulation of embryonic body size in the case of *M280;* Fig. 41C). Thus they are developmentally tolerated in the sense that they attract metabolically expensive pro-transcriptional machinery – p300 for sure and likely a menagerie of TFs – and yet the result of this pro-transcription milieu, which is gene transcription in the limb bud, is not even necessary for limb elaboration. Perhaps these enhancer sequences are highly conserved because of the physiological processes they regulate in adult life rather than during development.

One way to test this explanation in a laboratory setting would be to subject control and mutant *M280* and *M1442* adults to assays that require physical exertion that may mimic that experienced in a natural setting. For example, *myogenin* null mice exhibit deficiencies in skeletal muscle metabolism and exercise endurance only after run on a treadmill (Flynn et al., 2010; Meadows et al., 2011). Mice lacking *M280* or *M1442* could be assayed in a similar manner, perhaps with a focus on looking at the responses of limb elements to prolonged physical exertion.

In the context of the foregoing discussion it is worth noting that *M280* mice did exhibit a body weight phenotype, and this only strongly confirmed after birth. Six-week old male and female $M280^{tm1}$ -/- mice were on average about two grams lighter than control littermates. Both the developmental analyses and adult BMD/microCT scan results suggest that this phenotype is unrelated to the limbs, though this cannot be ruled out. Measurements of CR-length at E15.5 also showed that *M280* mutant embryos were smaller than control littermates. This observation

too appeared to be unrelated to limb development. Importantly, therefore, the phenotype that results from the genomic deletion of *M280* is manifested clearly in postnatal mice, not embryonic limb buds (Fig. 41C). This concrete observation suggests that *M280* is developmentally tolerated in the limb bud and that its extreme sequence conservation may be a result of selection pressures acting on its ability to regulate whole embryo body size and/or adult body weight. This demonstrates that an enhancer may have essential regulatory capacities not directly related to its entire spatial regulatory potential (Fig. 41C). Also, the adult expression patterns directed by either *M280* or *M1442* have not been documented, but certainly could be using the *M280*GWtg and *M1442*GWtg lines. Such documentation may provide insight into the roles these enhancers may play during mouse adult life or indicate potential sights of postnatal phenotypes that have gone undetected.

Explanations for the observed results have, to this point, focused on the action and function of enhancers – they may regulate gene transcription during embryonic or postnatal development and apparently may act in partially redundant fashion. It is also possible, however, that the unaffected limb development in mutant mice is a result of redundancy in genetic pathway activity rather than regulatory activity. For example, although *M1442* may be required for the proper limb expression of a particular gene, the effects of reduced transcription of that gene in mutant mice may be compensated for by another redundantly expressed gene that feeds into the same genetic circuit. Redundant gene function has been amply demonstrated for numerous gene families during limb development, including *Hoxa* and *Hoxd* cluster members, *Bmps* and *Fgfs (Niswander et al., 1993; Bandyopadhyay et al., 2006; Montavon et al., 2011)*. Thus, if *M280* or *M1442* regulates members of these genetic pathways and circuits then the potential phenotypic effects could be masked by redundant genetic inputs. But even this possibility does not satisfactorily explain why *M280* and *M1442* possess so many independent

signatures of enhancer activity and yet are not required in limb development (Bejerano et al., 2004; Visel et al., 2008; Visel et al., 2009a).

Genome-wide enhancer studies highlight potential causes of discrepancy between transgenic reporter assays and endogenous enhancer function

Further reason to suspect that transgenic experiments may not faithfully represent an enhancer's endogenous function come from the complex picture of enhancer behaviors drawn from experiments in cell lines, stem cell and otherwise. For example, Whyte, A. *et al.* (2013) demonstrated that large modular enhancers (median size approximately equal to 8500 bps), termed super-enhancers, are occupied by high concentrations of the TFs Klf4 and Esrrb and interact with high concentrations of Mediator (Whyte et al., 2013). Super-enhancers are particularly important in driving the transcription of genes necessary for stem cell pluripotency. Interestingly, depletion of either Mediator or Oct4 (also an important super-enhancer-bound TF) leads to preferential transcriptional disruption of pro-pluripotency genes, indicating that enhancers that regulate specific cellular states may also have specific signatures of bound TFs and activator complexes; such signatures may or may not be re-created at sites of transgene insertion.

In another report it was shown that long-range interactions between enhancers and promoters spanned, on average, 120 kbps, and that only 7% of these long-range interactions were between an enhancer and its nearest genes (Sanyal et al., 2012). Such long-range interactions simply cannot be reproduced at the random insertions sites of transgenes, thus in many instances it is possible that enhancer behavior in the context of its host transgene is not a complete representation of its endogenous activity. Another study showed that posttranslational modifications of nucleosomes conducive to enhancer-promoter interactions occurred within

domains such that groups of enhancer and promoters were modified together (Shen et al., 2012). But in transgenic experiments, enhancers are likely far removed from the coordinately regulated domain within which their endogenous loci reside; therefore their activity is likely altered in the transgenic context.

The above specific examples, along with those covered in numerous reviews, indicate that enhancers rarely, if ever, act in isolation (Malik and Roeder, 2010; Ong and Corces, 2011; Spitz and Furlong, 2012). Current evidence suggests that modular enhancers are themselves embedded within larger regulatory modules called by various names: enhancer-promoter units (Shen et al., 2012), archipelagos (Montavon et al., 2011), holo-enhancers (Marinic et al., 2013), long-range interaction landscapes (Sanyal et al., 2012) and super-enhancers (Whyte et al., 2013), to name a few. These various enhancer and enhancer network types indicate that many, rather than one or a few, regulatory elements contribute to the transcription of individual genes and groups of genes. The concerted effects of all of these regulatory inputs likely results in partial redundancy and dampening of genomic perturbation (including gene targeting) and consequently phenotypic outcomes. Thus, given what is known about enhancers in limb development specifically (see discussion above), and in other tissues and cell lines generally, it is not unreasonable to hypothesize that some form of enhancer network compensation is absorbing the developmental costs of the deletion of *M1442* and *M280* in embryonic mutant mouse limbs.

Potential experimental approaches to further explore the role of *M1442* and *M280* during limb development

The absence of a limb phenotype in *M280* and *M1442* null mice is a record of negative data. It is possible that *M280* and/or *M1442* do adversely affect limb development in some way.

The assays employed to assess limb development may not have detected an existing phenotype. This admission adds context to several conventional experiments that could be carried out to further interrogate the regulatory roles of *M280* and *M1442*. For example, it has been demonstrated that *Dmap1* and *Nme7* transcription is reduced to approximately half that of control levels in *M280* and *M1442* null mice, respectively. Therefore, it may be possible to compromise the genetic background of developing mutant embryos by crossing $M280^{tm1}$ -/- or $M1442^{tm1}$ -/- mice to heterozygous *Dmap1* or *Nme7* mutant mice, respectively. Homozygous *Dmap1* null embryos die early in development(Mohan et al., 2011). Homozygous *Nme7* null mice manifest phenotypes associated with defunct ciliary motility, including *situs inversus* (Vogel *et al.,* 2010). By crossing heterozygous *Dmap1* or *Nme7* mice with *M280* or *M1442* null mice, respectively, it may be possible to induce a limb-specific phenotype that reflects a genetic interaction between the enhancers and the genes.

Another logical extension of this research would be to use whole genome transcriptomics to identify candidate loci regulated by *M280* and *M1442*. The local transcriptional impact of deleting *M280* and *M1442* using qPCR has been presented in this study. And though significant transcriptional differences were noted, none of the genes that are misregulated are immediately relatable to limb development, nor are they expressed in a finely delineated manner as assayed with *in situ* hybridization. Collection of whole genome transcript levels via RNA-seq in mutant and control limbs at E11.5 may permit a clearer picture of the loci that are strongly associated with either *M280* or *M1442*. (Also, though unrelated to the limb, it would be worthwhile to collect whole genome transcript levels at E15.5, or slightly earlier, to relate particular transcript levels with the mutant body size detected in *M280* mutant embryos).

Bioinformatic evidence presented herein suggests that two previously described enhancers contain a collection of TF binding sites that are also found in *M280*. Perhaps these putative enhancers, which together span roughly 3 kb, serve as a partially redundant enhancer

for *M280*. This proposed *M280*-shadow-enhancer could be cloned upstream of a minimal promoter and *lacZ* and used in transgenic mouse experiments to address whether or not it directs limb-specific expression.

Lastly, *M280* and *M1442* may be useful as tissue-specific drivers of Cre recombinase or other proteins capable of modifying or acting on the mouse genome. For example, *M1442* directs transcription in specific wrist bones including the hamate, triquetral and pisiform, and possibly the scaphoid as well. Using *M1442* to drive Cre recombinase to these specific bones in order to modify target loci could permit a better understanding of the molecular genetic mechanisms that direct wrist development. The genetic mechanisms underlying the diversity of wrist bone elements are particularly important from an evolutionary perspective. Transformations in the wrist accompanied important transitions during vertebrate evolution including the so-called fin-to-limb transition as well as the unique autopod adaptations specific to the primate lineage (Sanchez-Villagra and Menke, 2005; Shubin et al., 2006; Prabhakar et al., 2008; Shubin et al., 2009; Mitgutsch et al., 2012; Cotney et al., 2013).

Implications for future studies aimed at deleting enhancers *in situ*

Both *M1442* and *M280* were deleted because their human orhtologs, *hs1442* and *hs280*, respectively, were capable of driving limb-specific expression in transgenic mouse embyros and because they were affiliated with p300, a transcriptional activator. However, nothing was known about the gene or genes these enhancers regulate. Consequently, interpreting our results in the light of a specific signaling pathway or genetic circuit known to be active during limb development is not possible. Furthermore, though *M280* mice manifest a weight phenotype we cannot at this point correlate the phenotype with any specific gene or genetic pathway for the same reasons described above. Thus in future studies that employ gene targeting to delete or

otherwise manipulate enhancers it will be important to rely on as many data sets as possible to select enhancers whose manipulation can advance our knowledge of enhancer behavior and development of specific structures the most. For example, in combination with the p300-ChIP-seq data set one could utilize knowledge of enhancer-gene proximity, published enhancer-promoter units, published catalogs of pro-transcription chromatin marks in specific tissues or cell types, and even data sets that take into account signatures of selection in the genome. The combination of such information should better lead researchers to enhancers that will likely yield tractable, mechanistic information upon endogenous manipulation.

It is important to re-emphasize the patience and attention to detail that ventures into enhancer function require. If indeed partial redundancy and robustness are persistent features of enhancer regulation then researchers need to be prepared to identify and characterize subtle phenotypes that result from functional experiments at endogenous enhancer loci (subtle relative to particular gene knockout experiments). Additionally, as these results indicate, it is imperative to remember that results of a particular enhancer-driven transgenic reporter assay do not necessarily predict the tissues and processes that will be affected upon deletion of that particular enhancer from the genome. Lastly, as enhancers continue to be deleted from the mouse genome it may be wise to consider experiments that foster conditions similar to those experienced by natural populations; it is in these contexts, especially after birth, that phenotypic consequences of enhancer deletions may be brought into sharper focus.

APPENDIX 1

Quantification of Transcript Levels in E11.5 fore- and hindlimb buds, Relative to Wild-type

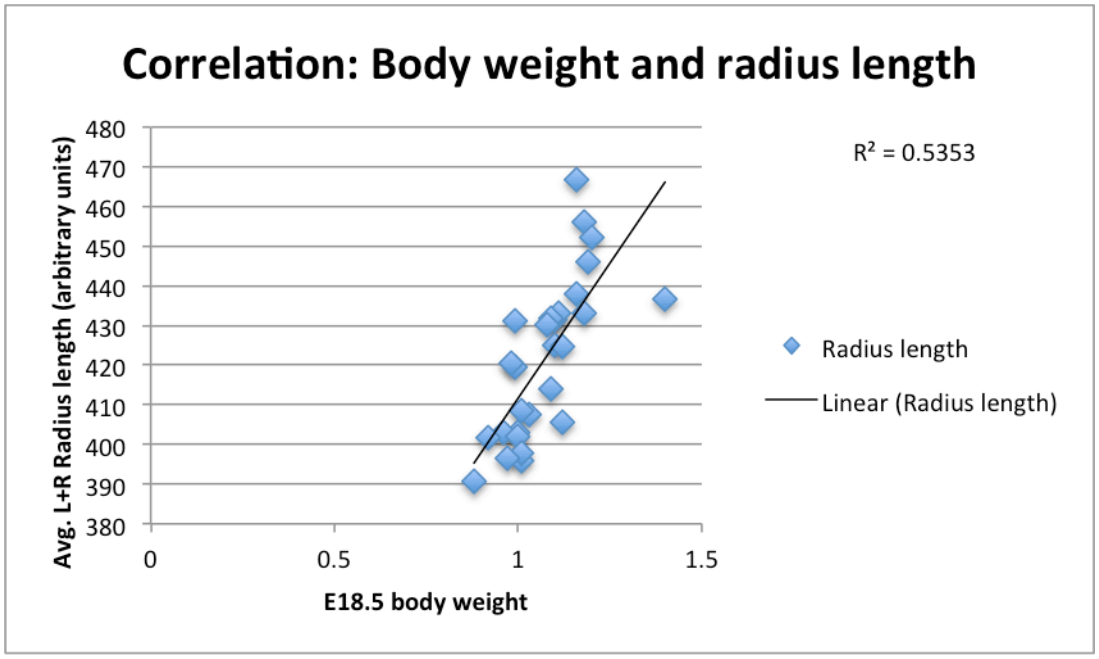
Numbers indicate average levels ± SD; each sample comprised of limb buds from three individuals (n=8)

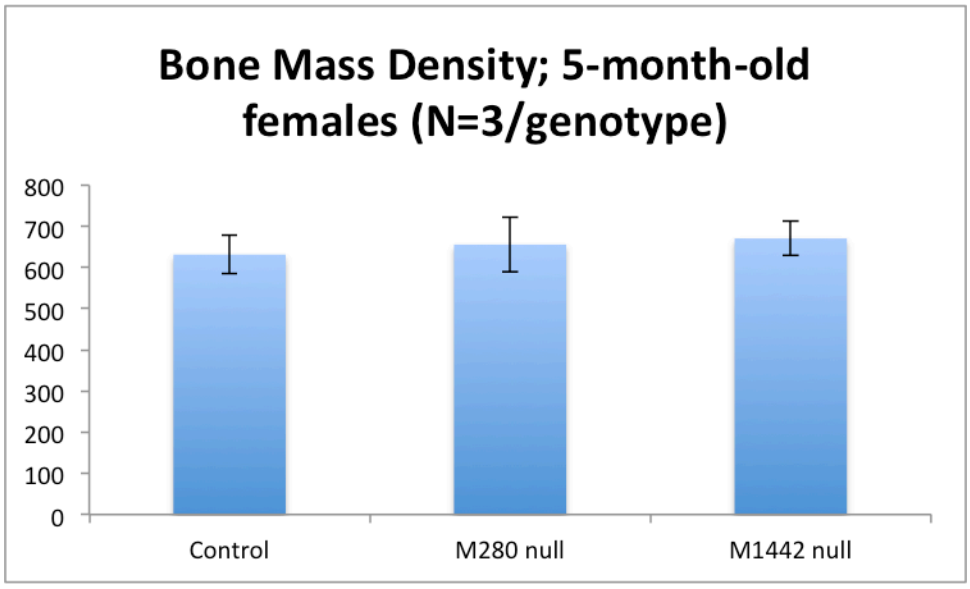
	Prdx1	*Tmem53*	*Rnf220*	*Eri3*	*Dmap1*
wt: *M280* +/+	1	1	1	1	1
M280 -/-	0.84 ± 0.10	0.62 ± 0.12	1.09 ± 0.09	0.79 ± 0.14	0.54 ± 0.15
P ≤ 0.05		✓			✓

Quantification of Transcript Levels in E11.5 fore- and hindlimb buds, Relative to Wild-type

Numbers indicate average levels ± SD; each sample comprised of limb buds from three individuals (n=8)

	Prx1	*Gorab*	*Kifap3*	*Scyl3*	*BC055324*	*Nme7*	*Atp1b1*
wt: *M1442* +/+	1	1	1	1	1	1	1
M1442 -/-	1.07 ± 0.18	1.31 ± 0.27	1.33 ± 0.15	1.34 ± 0.20	0.90 ± 0.20	0.65 ± 0.17	1.16 ± 0.07
P ≤ 0.05		✓	✓	✓		✓	


Correlation: Body weight and radius length
Avg. L+R Radius length (arbitrary units)
480
470
460
450
440
430
420
410
400
390
380
0
0.5
1
1.5
E18.5 body weight
R² = 0.5353
Radius length
Linear (Radius length)


Bone Mass Density; 5-month-old females (N=3/genotype)
800
700
600
500
400
300
200
100
0
Control
M280 null
M1442 null

Key: Matrices used to account for between-litter variation in CR-length in E15.5 embryos

Litter	ID	Genot.	C-R length
Color coded by litter	N = null N2 N3 N4		
	W = wt W2 W3 W4 W5		

NXN	1	2	3	4
1		(N1-N2)^2	etc.	etc.
2			(N2-N3)^2	etc.
3				etc.
4				

WXW	1	2	3	4	5
1		(W1-W2)^2	etc.	etc.	etc.
2			(W2-W3)^2	etc.	etc.
3				etc.	etc.
4					etc.
5					

NXW	1nl	2nl	3nl	4nl
1wt	(N1-W1)^2	etc.	etc.	etc.
2wt	etc.	etc.	etc.	etc.
3wt	etc.	etc.	etc.	etc.
4wt	etc.	etc.	etc.	etc.
5wt	etc.	etc.	etc.	etc.

#	N X N	W X W	N X W
1			
2			
3	N = 6		
4			
5		N = 10	
6			
7			
8			
9			
10			N = 20
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
Avg.			
S.D.			

Matrices: E15.5 CR-length (arbitrary units; Photoshop rule tool)

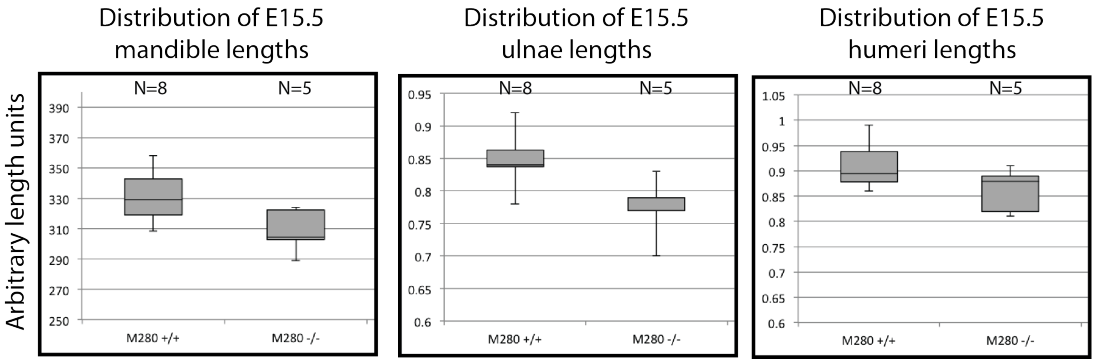
Litter	ID		Genot.	C-R length
	469	1	280 nu	981
	473	2	280 nu	989.6
	531	3	280 nu	966.6
	573	4	280 nu	895.7
	472		wt	1087
	534		wt	1039
	536		wt	1068
	567		wt	1043
	571		wt	1071

N X N	1	2	3	4
1		73.96	207.4	7276
2			529	8817
3				5027
4				

W X W	1	2	3	4	5
1		2275	353.4	1962	240.3
2			835.2	11.56	1037
3				650.3	10.89
4					829.4
5					

W X N	1nl	2nl	3nl	4nl
1wt	11215	9467	14472	36557
2wt	3387	2460	5271	20592
3wt	7586	6162	10302	29722
4wt	3795	2809	5776	21580
5wt	8172	6691	10983	30870

Sam.	N X N	W X W	W X N
1	73.96	2275	11215
2	207.4	353.4	9467
3	7276	1962	14472
4	529	240.3	36557
5	8817	835.2	3387
6	5027	11.56	2460
7		1037	5271
8		650.3	20592
9		10.89	7586
10		829.4	6162
11			10302
12			29722
13			3795
14			2809
15			5776
16			21580
17			8172
18			6691
19			10983
20			30870
Avg.	3655	820.6	12394
S.Dev	3902	771.8	10131


Distribution of E15.5 mandible lengths
Distribution of E15.5 ulnae lengths
Distribution of E15.5 humeri lengths
Arbitrary length units
N=8
N=5
390
370
350
330
310
290
270
250
M280 +/+
M280 -/-
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
1.05
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6

Transcription Factor Binding Sites located within *M280:* Used to find putative *M280* Shadow Enhancers	
Large Block (520 bp): defined by mouse-to-frog conservation via NCBI:Dcode (ECR Browser)	
AP2alphal/Tfap2a	a)Eckert, D. et al. 2005. The AP-2 family of transcription factors. Genome Bio. 6:246; b) Feng, W. et al. 2008. Identification and analysis of a conserved Tcfap2a intronic enhancer element required for expression in facial and limb bud mesenchyme. Mol Cell Biol. 28:315
E2F	a) Iaquinta, P. and Lees, J. 2007. Life and death decisions by the E2-F transcription factors. Curr Opin Cell Biol. 19:649; b) Cobrinik, D. et al. 1996. Shared role of the pRB-related p130 and p107 proteins in limb development. Genes Dev. 10:1633
GATA	Karamboulas, K. et al. 2010. Regulation of BMP-dependent chondrogenesis in early limb mesenchyme by TGFbeta signals. J. Cell Science 123:2068
Foxm1	Abassi, A. et al. 2010. Human intronic enhancers control distinct sub-domains of *Gli3* expression during mouse CNS and limb development. BMC Dev Bio. 10:44
CLOX/CUX1	Lizarraga, G. et al. 2002. Studies on the role of Cux1 in regulation of the onset of joint formation in the developing limb. Dev Biol. 243:44
Small Block (148 bp): defined by mouse-to-frog conservation via NCBI:Dcode (ECR Browser)	
ETS	Iwamoto, M. et al. 2007. Transcription factor ERG and joint and articular cartilage formation during mouse limb and spine skeletogenesis. Dev Biol. 305:40
Pea3	a) Eloy-Trinquet, S. et al. 2009. FGF signaling components are associated with muscles and tendons during limb development. Dev Dynamics 238:1195; b) Arber, S. et al. 2000. ETS Gene *Er81* Controls the Formation of Functional Connections between Group Ia Sensory Afferents and Motor Neurons. Cell 101:485; c) Ladle, D. and Frank, E. 2002. The role of the ETS gene PEA3 in the development of motor and sensory neurons. Physiol. and Behavior 77:571
CHX10 and HoxA4	HoxA4 (similar to CHX10) sight could be bound by any number of hox genes.
NRF2	a) Hinoi, E. et al. 2007. Nuclear factor E2 p45-related factor 2 negatively regulates chondrogenesis. Bone 40:377; b) Hinoi, E. et al. 2006. Nrf2 negatively regulates osteoblast differentiation via interfering with Runx2-dependent transcriptional activation. J Biol Chem. 281:18015

APPENDIX 2

THE UNIVERSITY OF TEXAS
MD Anderson Cancer Center

Memorandum

Department of Biostatistics

Telephone: 713 563-4297
Fax: 713 563-4243

Date: October 18, 2012

To: Richard Behringer, Ph.D.,
Professor

Mark Nolte,
Graduate Research Assistant
Department of Genetics

From: Caimiao Wei, Ph.D.,
Principal Statistical Analyst

Michele Guindani, Ph.D.,
Assistant Professor

Department of Biostatistics

Subject: **Analysis of mouse weight data.**

Objectives

The objectives of the analysis are to examine whether the weights of the mouse are different between genotypes (wild type/control, heterozygous, and mutant) for male and female mice separately.

Description of the data set

Body weight at 3 weeks and 6 weeks were measured in mice with litter size ranges from 3 to 10.

Methods

Summary statistics

Mouse weights were summarized using the mean (s.d.) and median (range) by group, time point, and gender.

Comparing body weights among genotypes

We use linear mixed models to examine the differences in body weights among genotypes, adjusted for litter size (as continuous variable). The linear mixed models include fixed effect of

genotype and litter size, and a random effect of litter to account for the correlation between mice within the same litters. Pairwise comparisons between genotypes were carried out using the least square of the means with adjustment for multiple comparisons within model using Tukey's method. We used quantile-quantile plots to examined the normality assumption of the residuals of the linear models. All analyses were performed using SAS 9.3 and R.

Summary of main findings

1. **Comparing body weight differences at 3 week in female mice.**

   The differences among genotypes are statistically significant (type 3 test of fixed effects $p < 0.0001$), with the mutant mice being the lightest, followed by heterozygous mice. The wild-type mice are the heaviest. All pairwise comparisons between genotypes were statistically significant at the 0.05 significance level after Tukey-Kramer adjustment for multiple comparisons. The inverse association between litter size and body weight was marginally significant ($p=0.09$).

2. **Comparing body weight differences at 3 week in male mice.**

   The differences in body weights among genotypes were statistically significant (type 3 test of fixed effects $p = 0.003$). Mutant mice had the significantly lighter weights than heterozygous and wild-type mice. The differences in body weights between heterozygous and wild-type mice were not statistically significant. The inverse association between litter size and body weight was marginally significant ($p=0.09$).

3. **Comparing body weight differences at 6 week in female mice.**

   The differences among genotypes were statistically significant (type 3 test of fixed effects $p < 0.0001$), with the mutant mice being the lightest, followed by heterozygous mice. The wild-type mice are the heaviest. With Tukey-Kramer adjustment for multiple comparisons, the difference between heterozygous and mutant mice was marginally significant (adjusted $p=0.07$). The differences between heterozygous and wild-type mice and between mutant mice and wild-type mice were significant after Tukey-Kramer adjustment. The association between litter size and body weight was not significant ($p=0.8$).

4. **Comparing body weight differences at 6 week in male mice.**

   The differences among genotypes were statistically significant (type 3 test of fixed effects $p < 0.0001$), with the mutant mice being the lightest, followed by heterozygous mice. The wild-type mice are the heaviest. With Tukey-Kramer adjustment for multiple comparisons, the difference between heterozygous and mutant mice was significant (adjusted $p=0.01$). The difference between heterozygous and wild-type mice was not significant ($p=0.25$), the difference between mutant mice and wild-type mice was significant ($p=0.001$). The association between litter size and body weight was not significant ($p=0.7$).

Detail Results:

i. Comparing body weight differences between genotypes

1. Comparing body weight differences at 3 week in female mice.

The differences in body weights among genotypes were statistically significant (type 3 test of fixed effects p <0.0001), with the mutant mice being the lightest, followed by heterozygous mice. The wild-type mice are the heaviest. All pairwise comparisons between genotypes were statistically significant at the 0.05 significance level after Tukey-Kramer adjustment for multiple comparisons. The inverse association between litter size and body weight was marginally significant (p=0.09). The residuals based on the raw weight scale are normally distributed.

Solution for Fixed Effects

Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		11.7283	0.9842	18	11.92	<.0001
Genotype	het	-0.7192	0.2407	51	-2.99	0.0043
Genotype	null	-1.5255	0.2952	51	-5.17	<.0001
Genotype	wt	0	.	.	.	.
litter_size		-0.2454	0.1429	51	-1.72	0.0919

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Genotype	2	51	13.52	<.0001
litter_size	1	51	2.95	0.0919

Least Squares Means

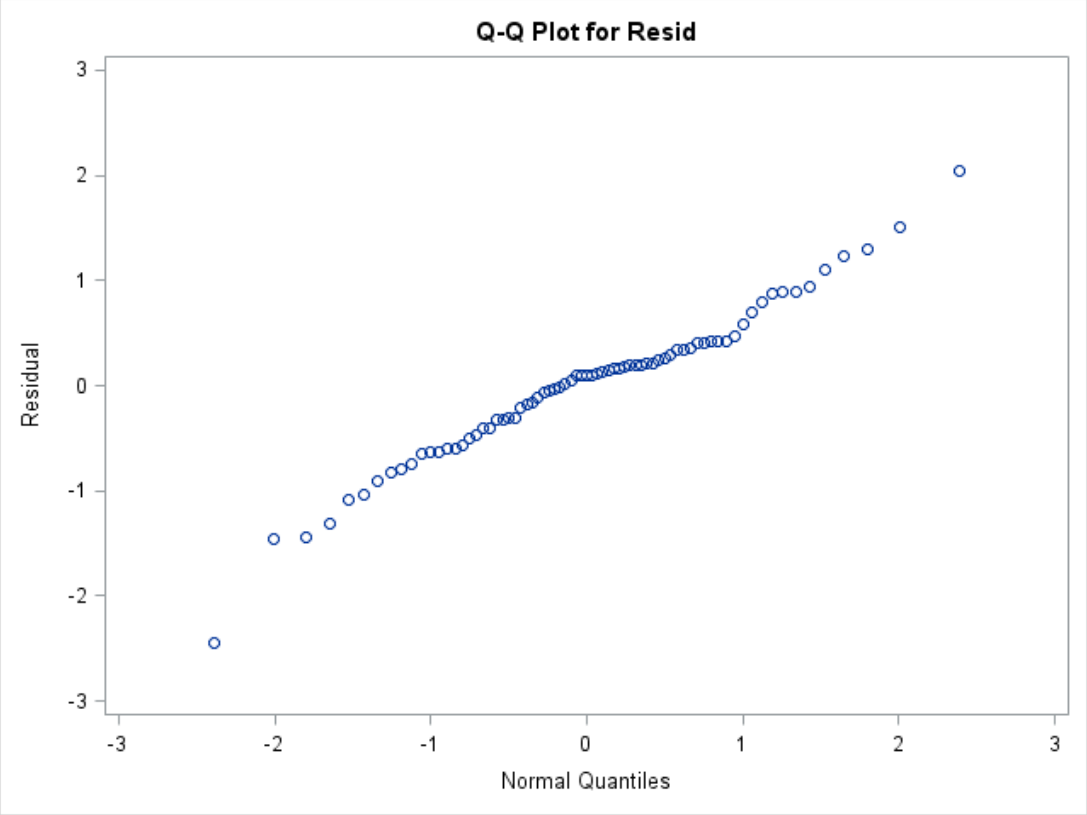
Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Genotype	het	9.3519	0.2478	51	37.73	<.0001
Genotype	null	8.5456	0.3035	51	28.16	<.0001
Genotype	wt	10.0711	0.2646	51	38.07	<.0001

Differences of Least Squares Means

Effect	Genotype	_Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P

Differences of Least Squares Means

Effect	Genotype	_Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
Genotype	**het**	**null**	0.8063	0.2773	51	2.91	0.0054	Tukey-Kramer	0.0146
Genotype	**het**	**wt**	-0.7192	0.2407	51	-2.99	0.0043	Tukey-Kramer	0.0118
Genotype	**null**	**wt**	-1.5255	0.2952	51	-5.17	<.0001	Tukey-Kramer	<.0001



2. Comparing body weight differences at 3 week in male mice.

The differences among genotypes are statistically significant (type 3 test of fixed effects p =0.003). Mutant mice had the significantly lighter weights than heterozygous and wild-type mice (adjusted p-values =0.04, and 0.003, respectively). The differences in body weights between heterozygous and wild-type mice were not statistically significant. The inverse association between litter size and body weight was marginally significant (p=0.09)..

Solution for Fixed Effects

Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		12.1065	1.2671	20	9.55	<.0001
Genotype	het	-0.4965	0.2966	55	-1.67	0.0997
Genotype	null	-1.0715	0.3111	55	-3.44	0.0011
Genotype	wt	0	.	.	.	.
litter_size		-0.3069	0.1784	55	-1.72	0.0909

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Genotype	2	55	6.61	0.0027
litter_size	1	55	2.96	0.0909

Least Squares Means

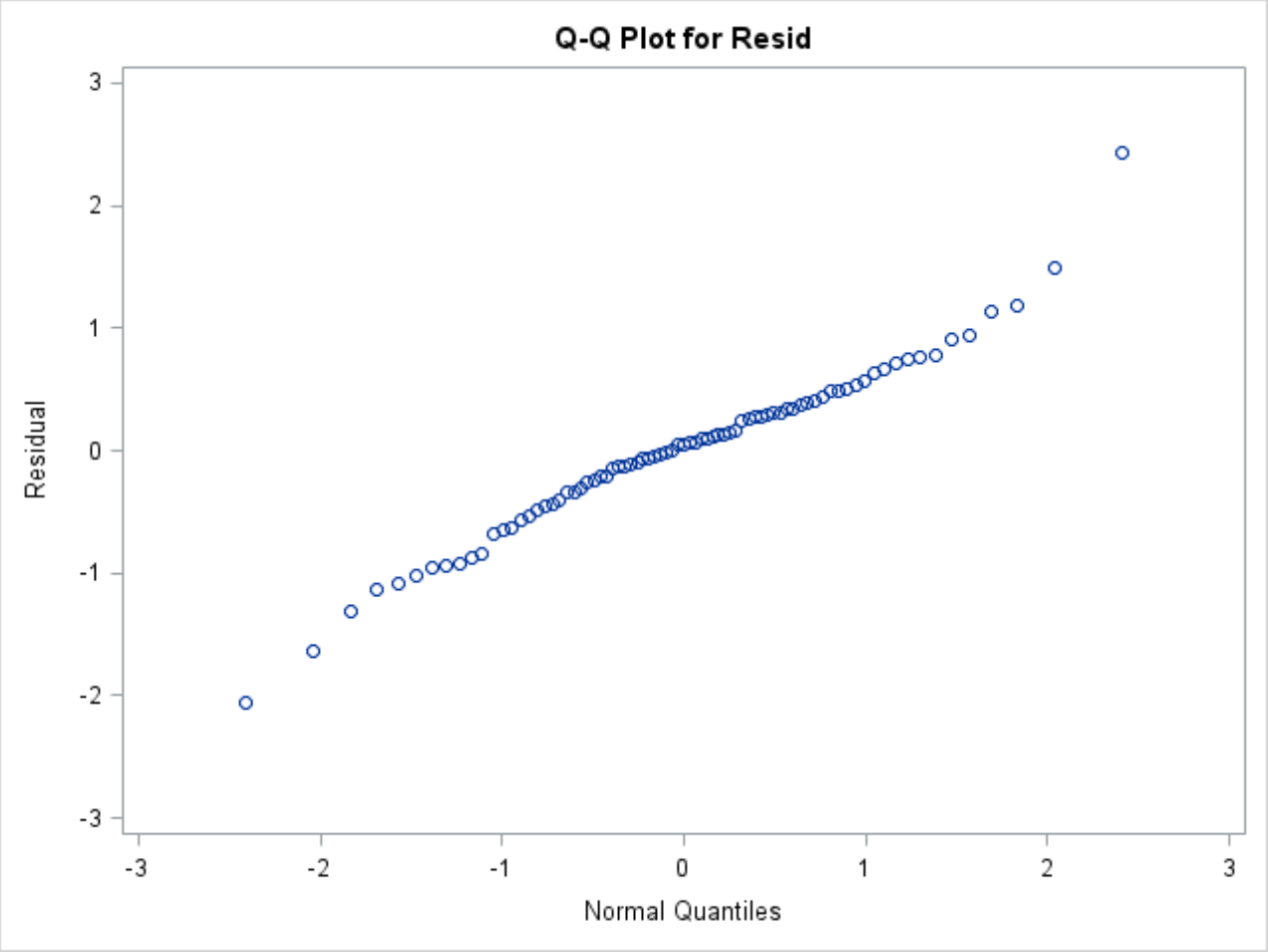
Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Genotype	het	9.4302	0.3032	55	31.10	<.0001
Genotype	null	8.8552	0.3294	55	26.89	<.0001
Genotype	wt	9.9268	0.3683	55	26.95	<.0001

Differences of Least Squares Means

Effect	Genotype	_Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
Genotype	het	null	0.5750	0.2274	55	2.53	0.0144	Tukey-Kramer	0.0376
Genotype	het	wt	-0.4965	0.2966	55	-1.67	0.0997	Tukey-Kramer	0.2241

Differences of Least Squares Means

Effect	Genotype	_Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
Genotype	**null**	**wt**	-1.0715	0.3111	55	-3.44	0.0011	Tukey-Kramer	0.0031



3. Comparing body weight differences at 6 week in female mice.

The differences among genotypes are statistically significant (type 3 test of fixed effects p <0.0001), with the mutant mice being the lightest, followed by heterozygous mice. The wild-type mice are the heaviest. With Tukey-Kramer adjustment for multiple comparisons, the difference between heterozygous and mutant mice was marginally significant (adjusted p=0.07). The differences between heterozygous and wild-type mice and between mutant mice and wild-type mice were significant after Tukey-Kramer adjustment (adjusted p-value=0.01, and p<0.0001, respectively). The association between litter size and body weight was not significant (p=0.8).

Solution for Fixed Effects

Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		17.3571	0.9954	18	17.44	<.0001
Genotype	het	-1.2724	0.4218	49	-3.02	0.0040
Genotype	null	-2.3136	0.4887	49	-4.73	<.0001
Genotype	wt	0	.	.	.	.
litter_size		0.02883	0.1413	49	0.20	0.8392

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Genotype	2	49	11.51	<.0001
litter_size	1	49	0.04	0.8392

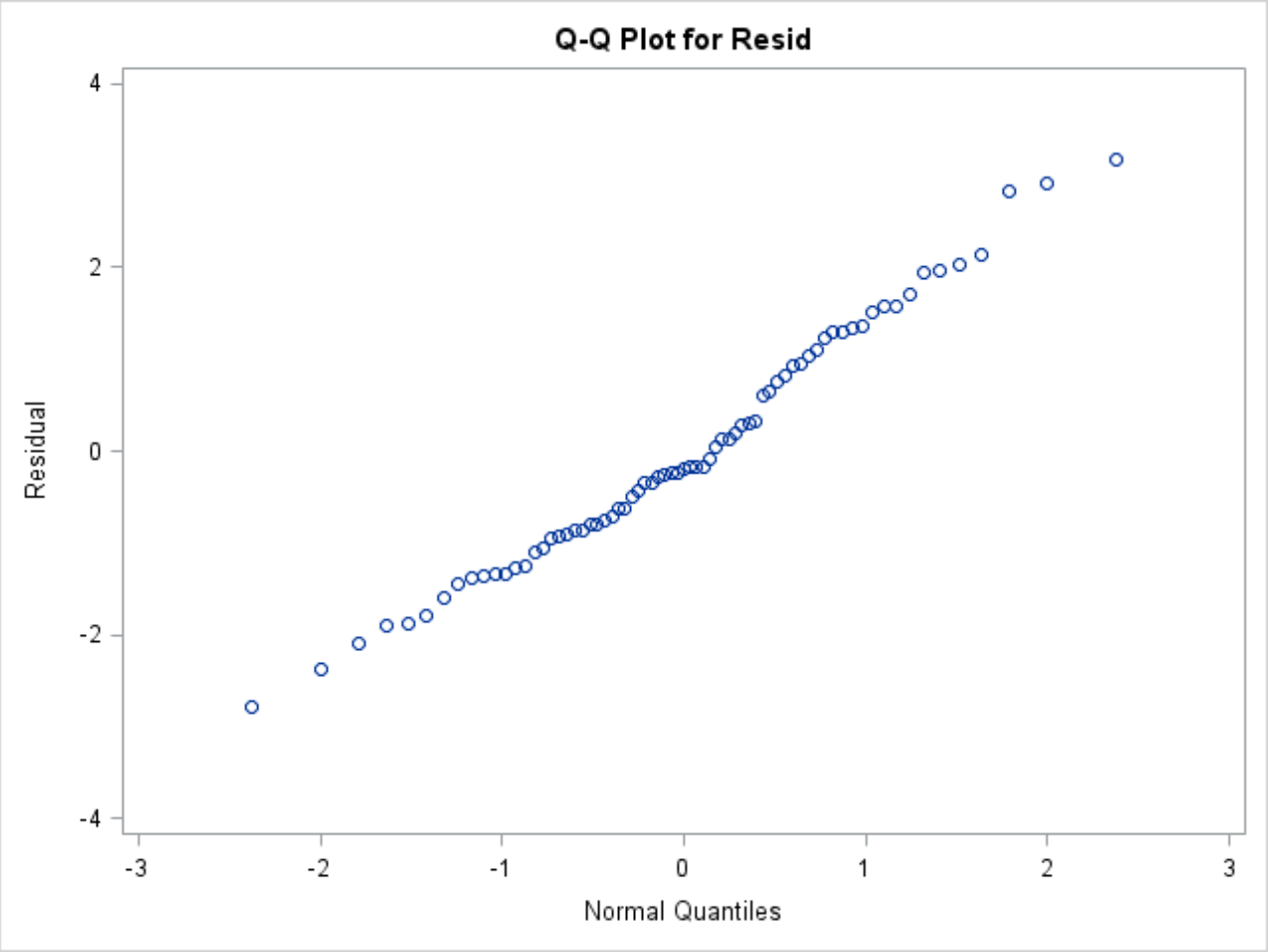
Least Squares Means

Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Genotype	het	16.2755	0.3259	49	49.93	<.0001
Genotype	null	15.2343	0.4088	49	37.26	<.0001
Genotype	wt	17.5479	0.3593	49	48.83	<.0001

Differences of Least Squares Means

Effect	Genotype	_Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
Genotyp	het	null	1.0412	0.4620	49	2.25	0.028	Tukey-	0.072

Differences of Least Squares Means									
Effect	**Genotype**	**_Genotype**	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**	**Adjustment**	**Adj P**
e							7	Kramer	2
Genotype	**het**	**wt**	-1.2724	0.4218	49	-3.02	0.0040	Tukey-Kramer	0.0111
Genotype	**null**	**wt**	-2.3136	0.4887	49	-4.73	<.0001	Tukey-Kramer	<.0001


Q-Q Plot for Resid
Residual
Normal Quantiles
4
2
0
-2
-4
-3
-2
-1
0
1
2
3

4. Comparing body weight differences at 6 week in male mice.

The differences among genotypes are statistically significant (type 3 test of fixed effects p <0.0001), with the mutant mice being the lightest, followed by heterozygous mice. The wild-type mice are the heaviest. With Tukey-Kramer adjustment for multiple comparisons, the difference between heterozygous and mutant mice was significant (adjusted p=0.01). The difference between heterozygous and wild-type mice was not significant (adjusted p=0.25), the difference between mutant mice and wild-type mice was significant (adjusted p=0.001). The association between litter size and body weight was not significant (p=0.7).

Solution for Fixed Effects

Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept		20.0895	1.3848	21	14.51	<.0001
Genotype	**het**	-0.9064	0.5576	54	-1.63	0.1099
Genotype	**null**	-2.3147	0.6179	54	-3.75	0.0004
Genotype	**wt**	0	.	.	.	.
litter_size		0.07039	0.1934	54	0.36	0.7173

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
Genotype	2	54	7.97	0.0009
litter_size	1	54	0.13	0.7173

Least Squares Means

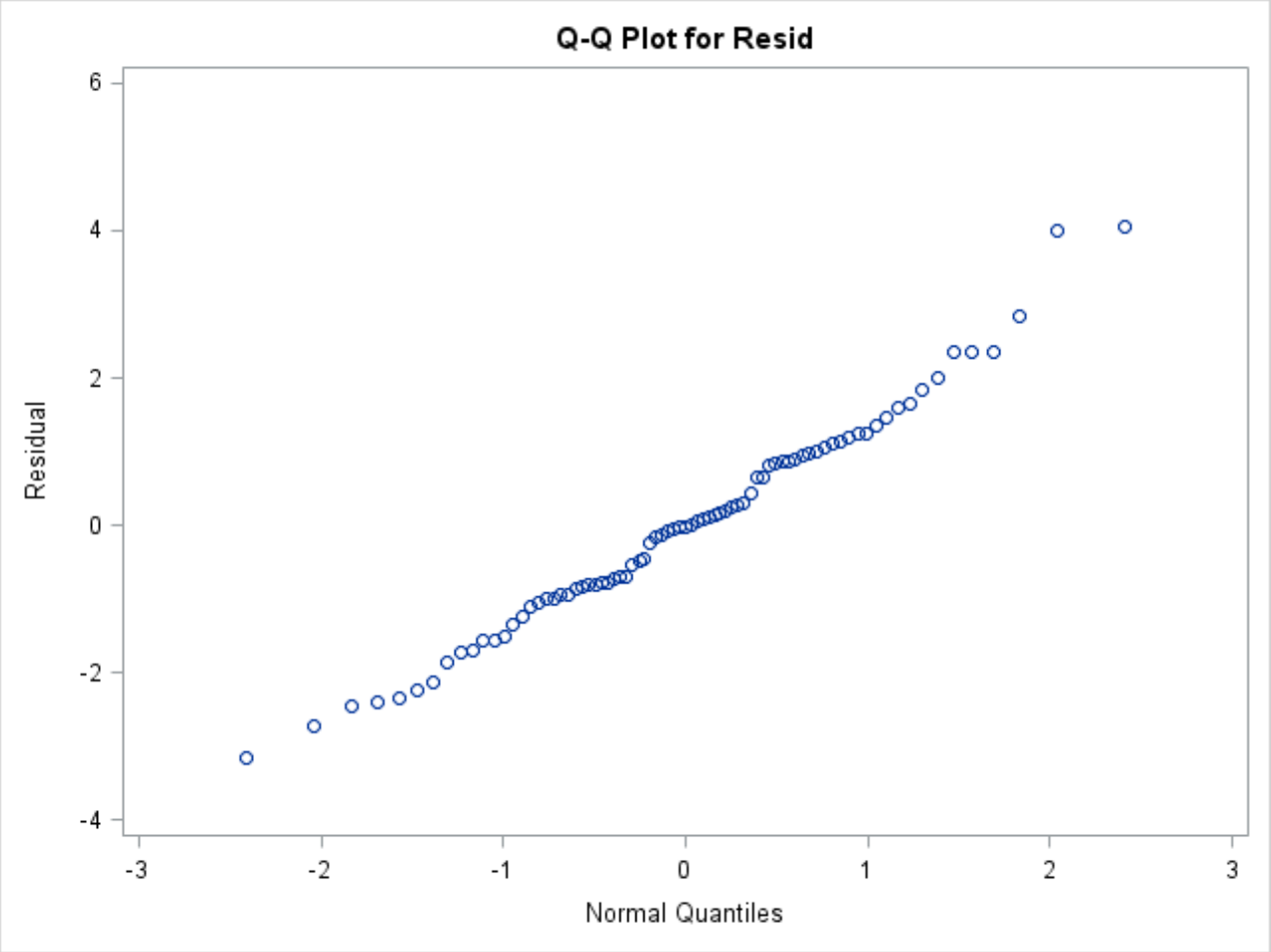
Effect	Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|
Genotype	**het**	19.6749	0.3972	54	49.53	<.0001
Genotype	**null**	18.2666	0.4904	54	37.25	<.0001
Genotype	**wt**	20.5813	0.5594	54	36.79	<.0001

Differences of Least Squares Means

Effect	Genotype	_Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
Genotype	**het**	**null**	1.4083	0.4605	54	3.06	0.0035	Tukey-Kramer	0.0095
Genotype	**het**	**wt**	-0.9064	0.5576	54	-1.63	0.109	Tukey-	0.243

Differences of Least Squares Means

Effect	Genotype	_Genotype	Estimate	Standard Error	DF	t Value	Pr > |t|	Adjustment	Adj P
e							9	Kramer	8
Genotype	**null**	**wt**	-2.3147	0.6179	54	-3.75	0.0004	Tukey-Kramer	0.0013



ii. Summary statistics

Table 1. Summary of mouse weights by gender, time point, litter size, and genotype.

At 3 week:

Gender	litter_size	Genotype	mean	std	max	median	min
f	3	null	.	.	.	.	.
f	3	wt	.	.	.	.	.
f	4	het	10.5500	3.18198	12.8	10.55	8.3
f	4	null	10.8000	.	10.8	10.80	10.8
f	4	wt	10.9000	1.83848	12.2	10.90	9.6
f	5	het	10.0000	0.42426	10.3	10.00	9.7
f	5	null	8.9500	0.63640	9.4	8.95	8.5
f	5	wt	11.1333	0.15275	11.3	11.10	11.0
f	6	het	9.1941	1.05087	11.4	9.20	7.1
f	6	null	8.7429	0.72078	9.8	9.00	7.8
f	6	wt	10.1600	0.82037	11.1	10.20	8.9
f	7	het	7.4500	2.19203	9.0	7.45	5.9
f	7	wt	9.7000	1.27671	11.1	9.40	8.6
f	8	het	10.0000	1.30894	11.3	10.50	7.5
f	8	null	9.1500	1.67033	10.9	9.35	7.0
f	8	wt	9.6857	1.06994	10.7	9.90	8.0
f	9	het	8.7000	0.96437	9.4	9.10	7.6
f	9	wt	10.4000	0.42426	10.7	10.40	10.1
f	10	het	9.0000	.	9.0	9.00	9.0
f	10	null	7.8000	.	7.8	7.80	7.8
f	10	wt	8.7000	0.84853	9.3	8.70	8.1
m	3	het	.	.	.	.	.
m	3	wt	.	.	.	.	.
m	4	het	10.7400	2.47952	13.4	10.70	7.8
m	4	null	13.4000	.	13.4	13.40	13.4

Gender	litter_size	Genotype	mean	std	max	median	min
m	5	het	11.6000	.	11.6	11.60	11.6
m	5	wt	9.9000	.	9.9	9.90	9.9
m	6	het	9.2692	1.46024	11.2	9.50	5.4
m	6	null	9.0375	1.29608	10.3	9.30	6.9
m	6	wt	10.7667	0.23094	10.9	10.90	10.5
m	7	het	9.3600	1.70822	11.7	9.70	7.2
m	7	null	8.2667	0.65064	8.9	8.30	7.6
m	7	wt	9.7000	.	9.7	9.70	9.7
m	8	het	9.4231	0.66226	10.6	9.40	8.3
m	8	null	9.1375	0.93799	10.3	8.95	7.6
m	8	wt	10.5571	0.83238	11.7	10.80	9.0
m	9	het	8.0333	1.41892	9.3	8.30	6.5
m	9	null	6.1000	.	6.1	6.10	6.1
m	10	het	8.7333	0.32146	9.1	8.60	8.5
m	10	null	7.9000	.	7.9	7.90	7.9
m	10	wt	9.0000	0.14142	9.1	9.00	8.9

At 6 weeks:

Gender	litter_size	Genotype	mean	std	max	median	min
f	3	null	15.5000	1.97990	16.9	15.50	14.1
f	3	wt	19.7000	.	19.7	19.70	19.7
f	4	het	15.8500	0.63640	16.3	15.85	15.4
f	4	null	14.8000	.	14.8	14.80	14.8
f	4	wt	16.1000	1.55563	17.2	16.10	15.0
f	5	het	16.5000	1.41421	17.5	16.50	15.5
f	5	null	15.4000	3.67696	18.0	15.40	12.8
f	5	wt	18.9667	1.46401	20.3	19.20	17.4
f	6	het	15.9533	1.64006	19.4	15.40	13.7
f	6	null	15.0000	1.26227	16.9	14.70	13.2
f	6	wt	17.2250	1.93800	19.8	16.90	15.3

Gender	litter_size	Genotype	mean	std	max	median	min
f	7	het	13.2000	.	13.2	13.20	13.2
f	7	wt	18.1500	2.75772	20.1	18.15	16.2
f	8	het	17.0857	1.20752	18.5	17.80	15.6
f	8	null	15.7500	2.25758	19.1	14.85	14.2
f	8	wt	16.6714	1.41270	18.9	16.20	14.9
f	9	het	15.7667	1.40119	16.9	16.20	14.2
f	9	wt	18.2500	0.77782	18.8	18.25	17.7
f	10	het	18.4000	.	18.4	18.40	18.4
f	10	null	15.6000	.	15.6	15.60	15.6
f	10	wt	18.1500	0.35355	18.4	18.15	17.9
m	3	het	18.6000	.	18.6	18.60	18.6
m	3	wt	21.4500	0.35355	21.7	21.45	21.2
m	4	het	19.8400	2.99299	24.0	18.70	17.1
m	4	null	20.9000	.	20.9	20.90	20.9
m	5	het	19.6000	.	19.6	19.60	19.6
m	5	wt	20.9000	.	20.9	20.90	20.9
m	6	het	19.0231	2.46751	22.9	18.80	15.3
m	6	null	18.1333	2.47925	23.0	16.90	16.6
m	6	wt	20.9500	0.63640	21.4	20.95	20.5
m	7	het	20.6200	2.48737	23.7	21.20	17.7
m	7	null	18.6333	2.17792	20.4	19.30	16.2
m	7	wt	21.7000	.	21.7	21.70	21.7
m	8	het	19.6769	1.77208	22.1	19.60	16.3
m	8	null	18.5000	2.41306	23.0	18.10	15.8
m	8	wt	20.8714	1.39847	23.1	20.80	18.5
m	9	het	19.5000	1.73494	21.4	19.10	18.0
m	9	null	15.2000	.	15.2	15.20	15.2
m	10	het	21.3667	0.58595	21.8	21.60	20.7
m	10	null	19.5000	.	19.5	19.50	19.5
m	10	wt	22.0000	1.55563	23.1	22.00	20.9

Figure 1. Mouse weights at 3 week by genotype and gender averaged over litter size.

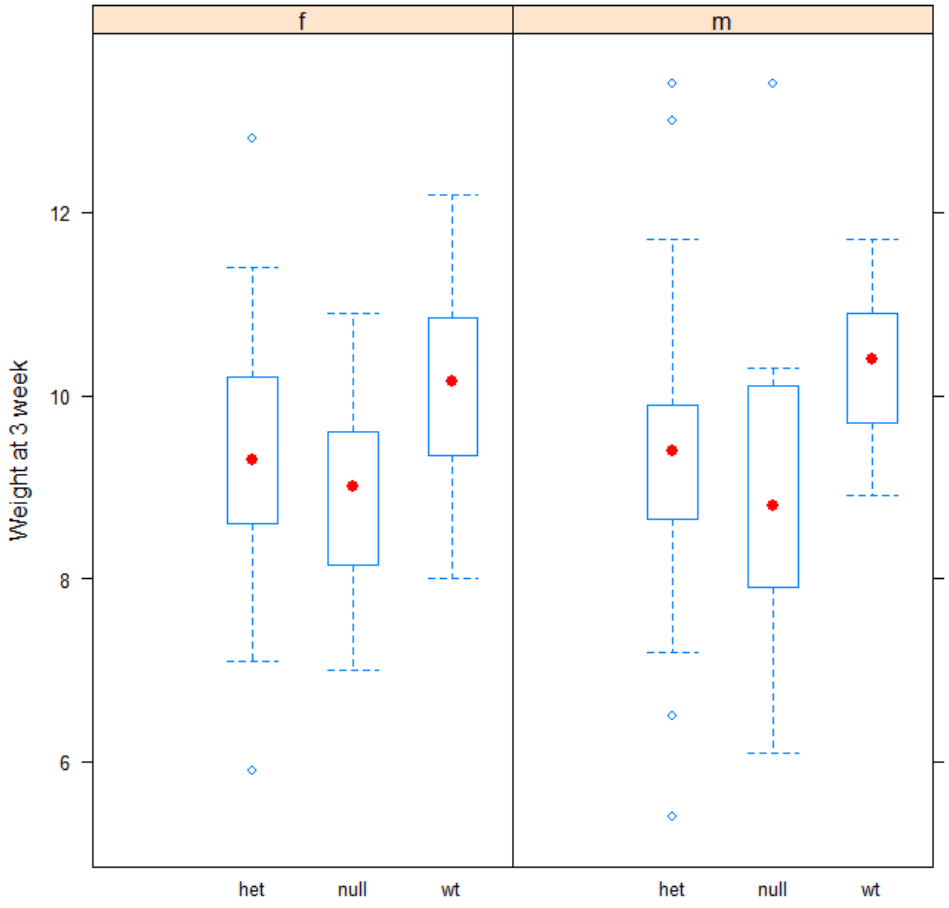


Figure 2. Mouse weights at 6 week by genotype and gender averaged over litter size.

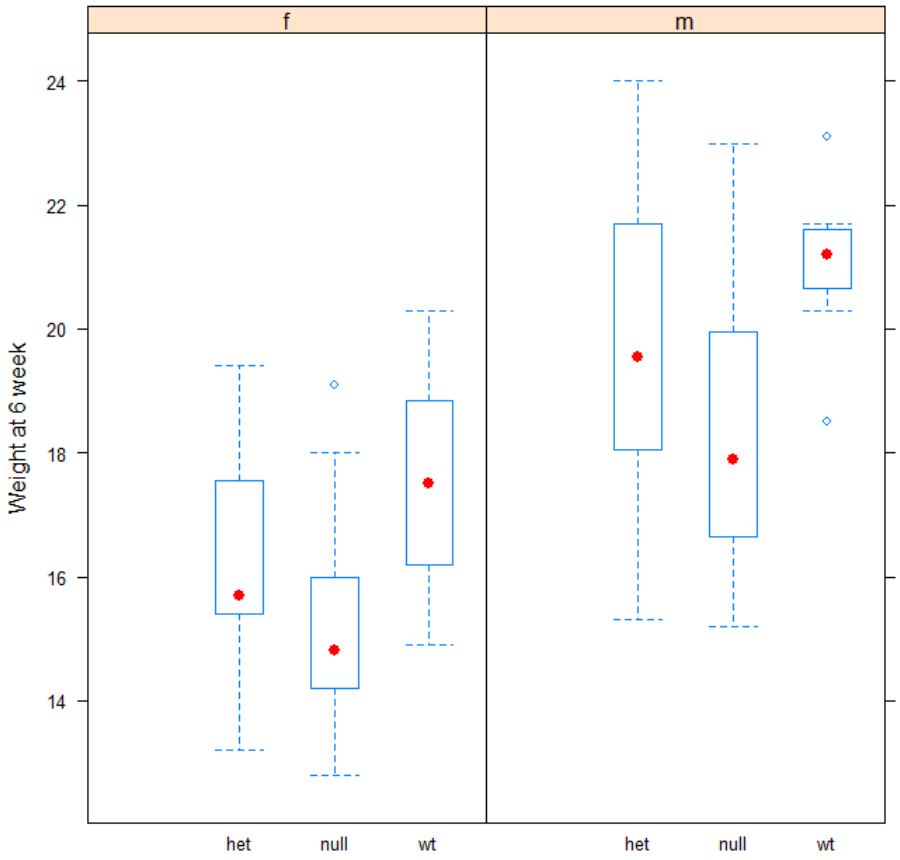


Figure 3. Female mouse weights at 3 week by litter size with genotype (for trend in litter size).

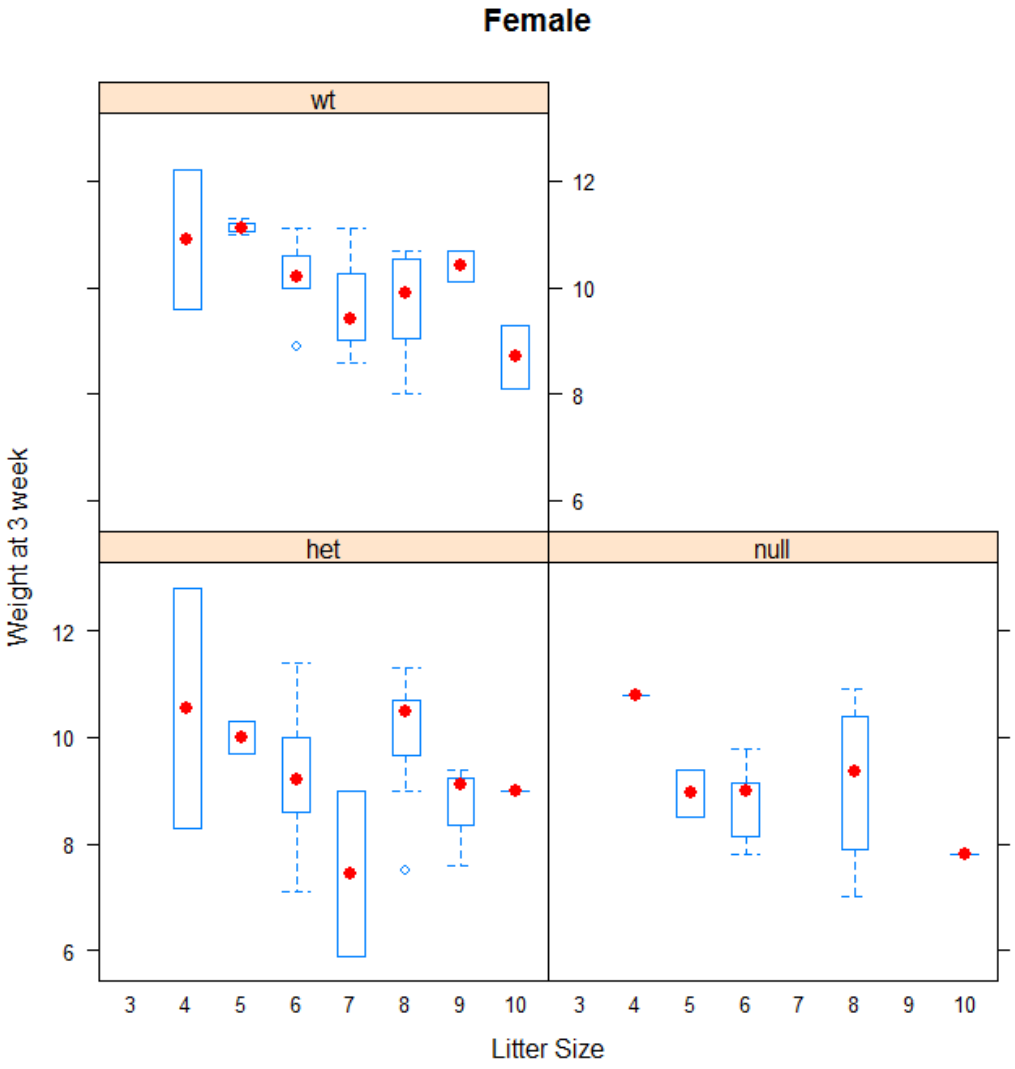


Figure 4. Female mouse weights at 3 week by litter size and genotype

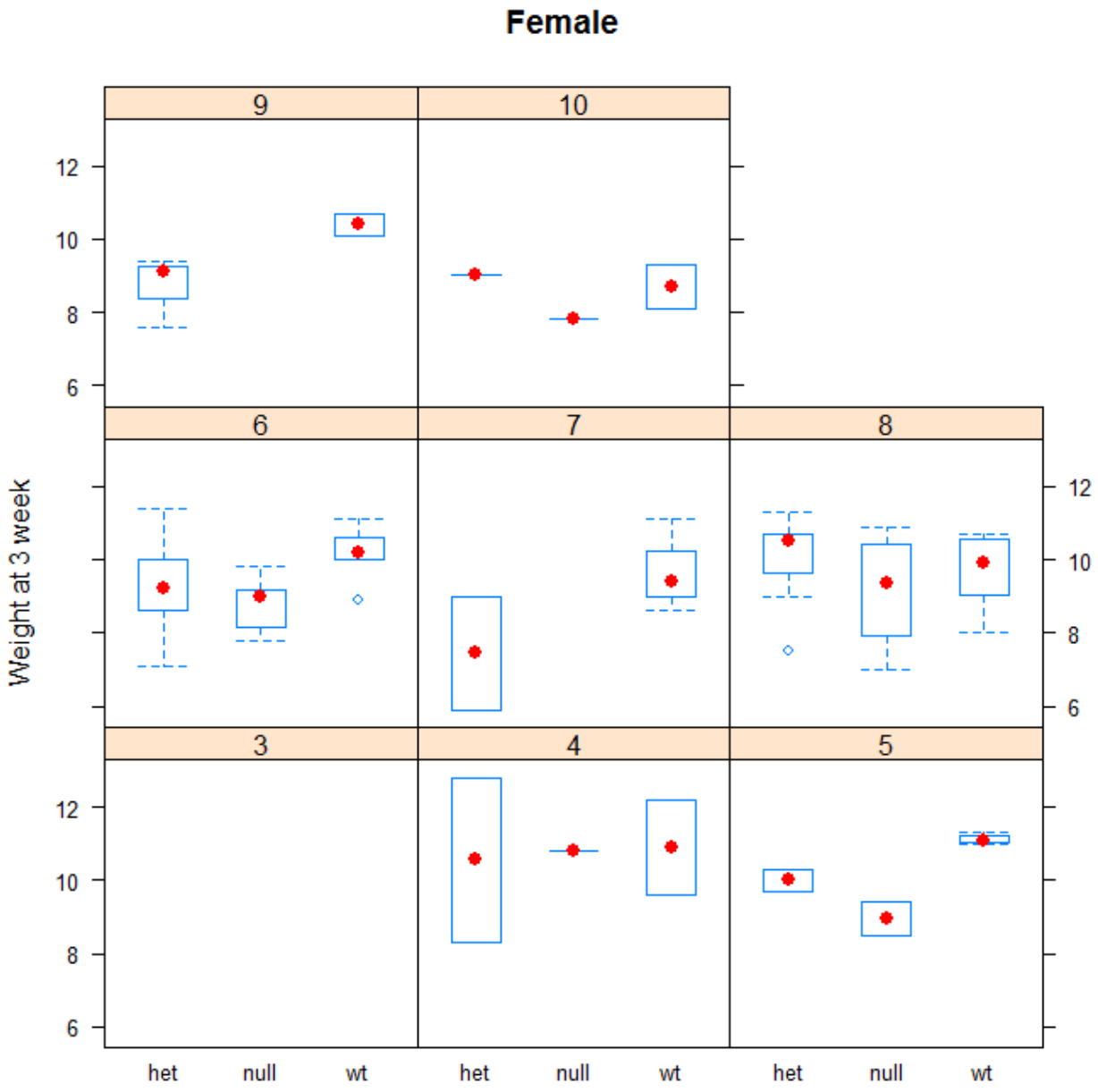


Figure 5. Female mouse weights at 6 week by litter size with genotype (for trend in litter size).

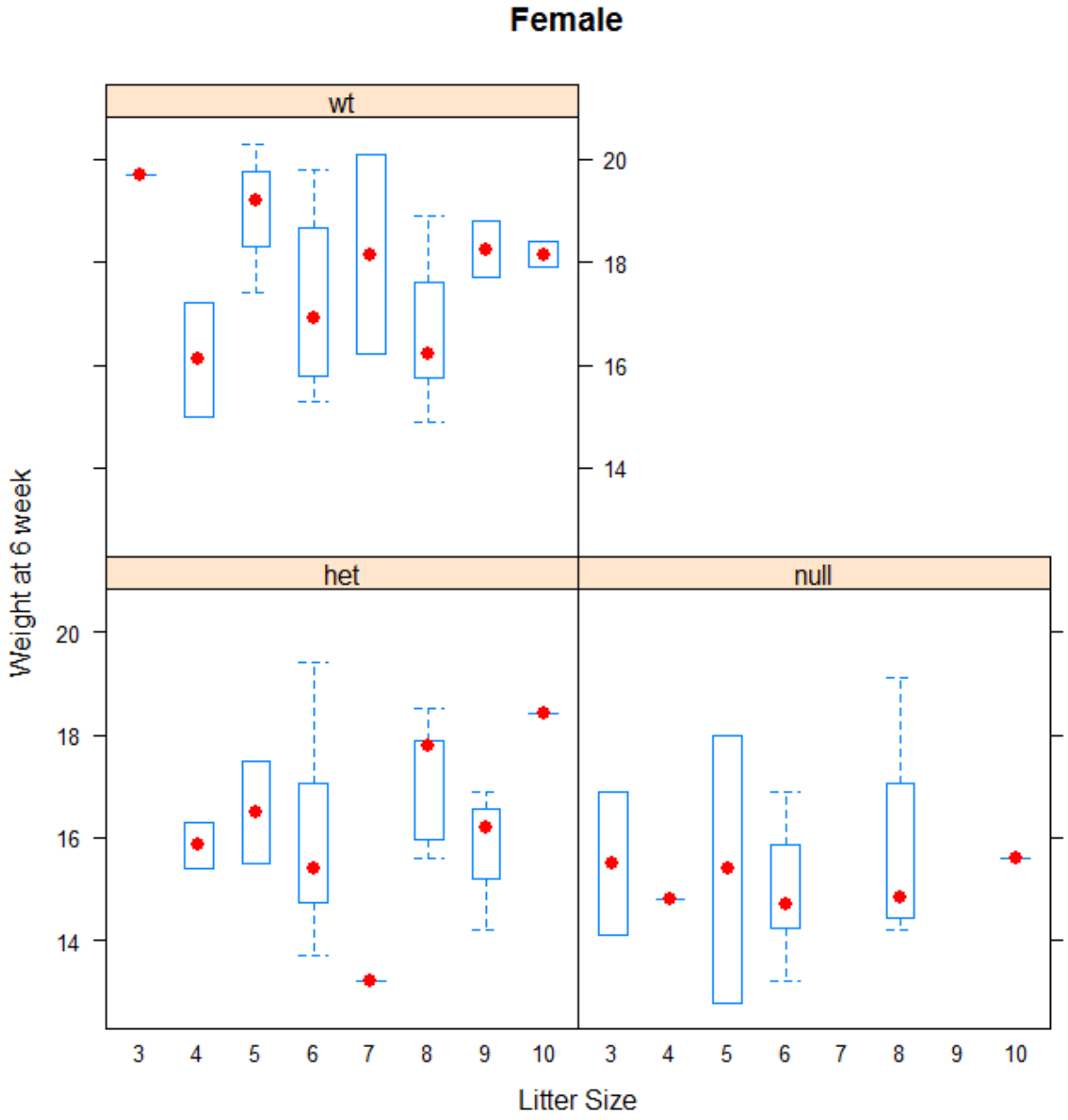


Figure 6. Female mouse weights at 6 week by litter size and genotype.

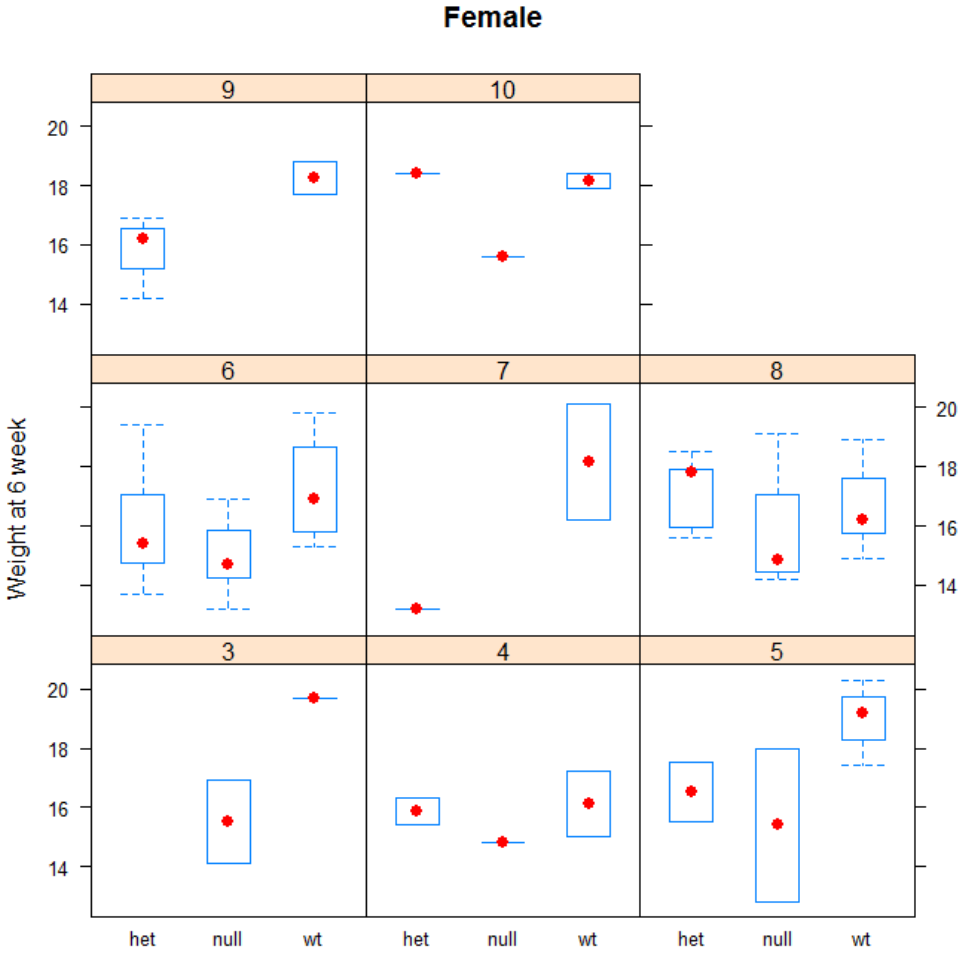


Figure 7. Male mouse weights at 3 week by litter size with genotype (for trend in litter size).

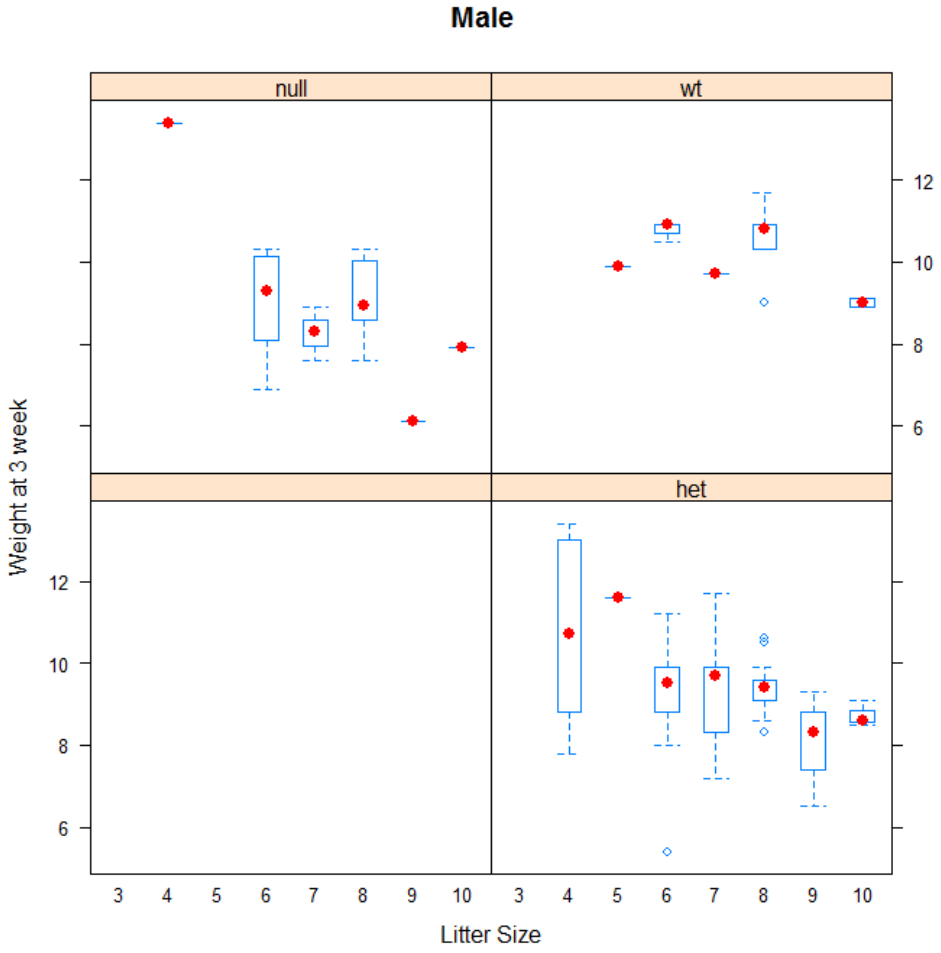


Figure 8. Male mouse weights at 3 week by litter size and genotype

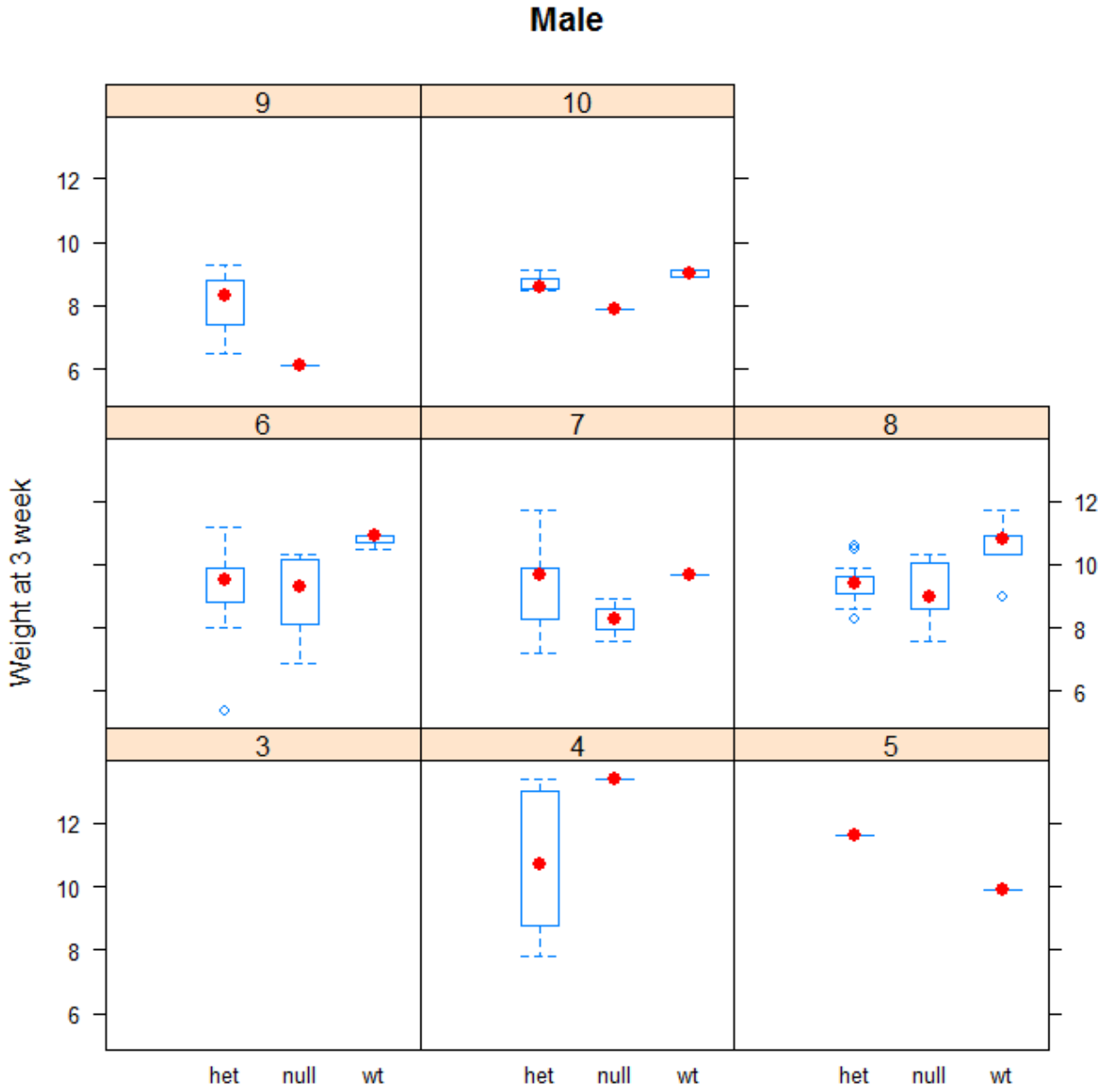


Figure 9. Male mouse weights at 6 week by litter size with genotype (for trend in litter size).

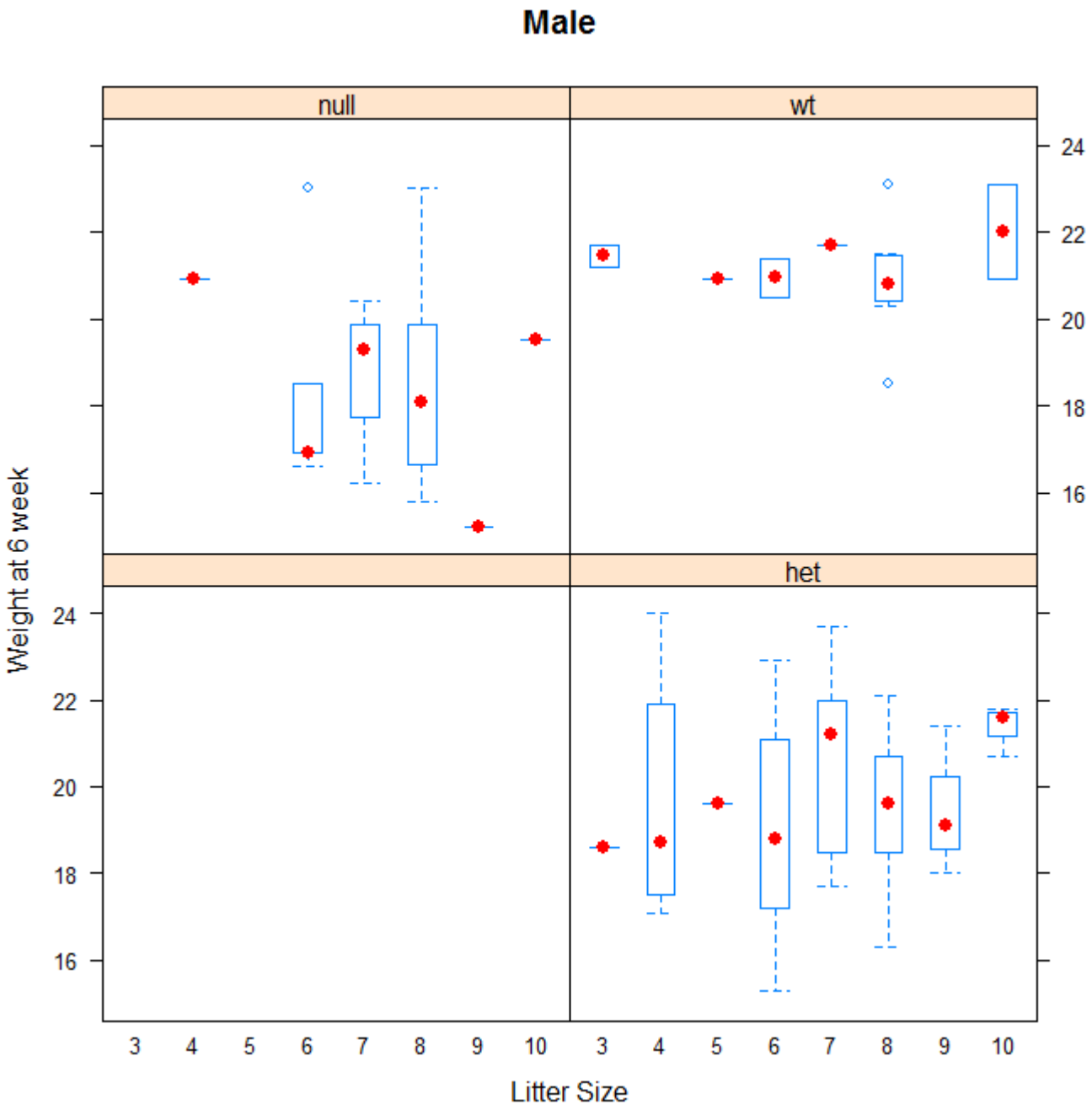
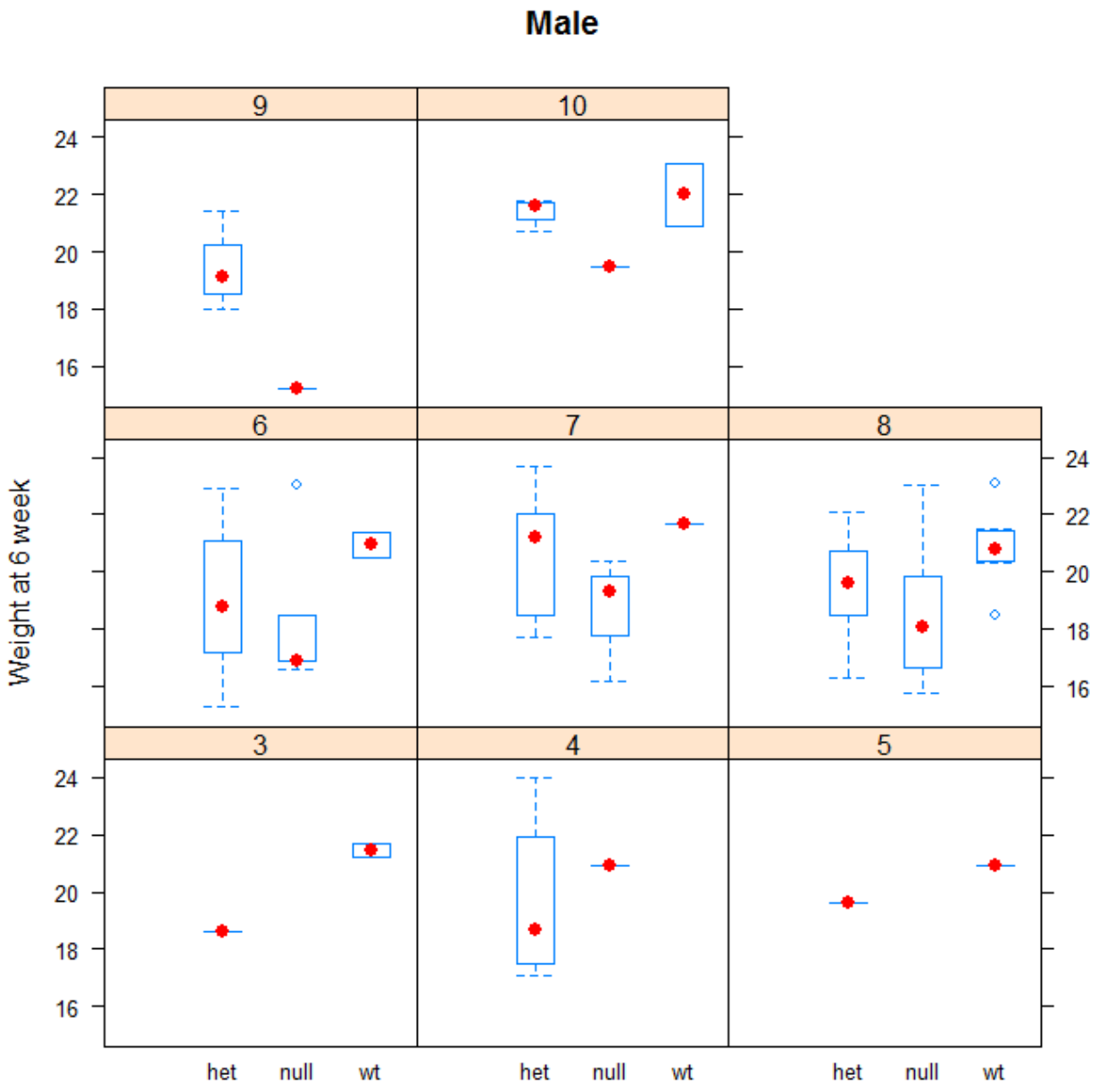


Figure 10. Male mouse weights at 6 week by litter size and genotype.



BIBLIOGRAPHY

Ahituv, N., Prabhakar, S., Poulin, F., Rubin, E. M. and Couronne, O. (2005) 'Mapping cis-regulatory domains in the human genome using multi-species conservation of synteny', *Hum Mol Genet* 14(20): 3057-63.

Ahituv, N., Zhu, Y., Visel, A., Holt, A., Afzal, V., Pennacchio, L. A. and Rubin, E. M. (2007) 'Deletion of ultraconserved elements yields viable mice', *PLoS Biol* 5(9): e234.

Baguna, J. and Garcia-Fernandez, J. (2003) 'Evo-Devo: the long and winding road', *Int J Dev Biol* 47(7-8): 705-13.

Bandyopadhyay, A., Tsuji, K., Cox, K., Harfe, B. D., Rosen, V. and Tabin, C. J. (2006) 'Genetic analysis of the roles of BMP2, BMP4, and BMP7 in limb patterning and skeletogenesis', *PLoS Genet* 2(12): e216.

Banerji, J., Rusconi, S. and Schaffner, W. (1981) 'Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences', *Cell* 27(2 Pt 1): 299-308.

Behringer, R. R., Finegold, M. J. and Cate, R. L. (1994) 'Mullerian-inhibiting substance function during mammalian sexual development', *Cell* 79(3): 415-25.

Bejerano, G., Pheasant, M., Makunin, I., Stephen, S., Kent, W. J., Mattick, J. S. and Haussler, D. (2004) 'Ultraconserved elements in the human genome', *Science* 304(5675): 1321-5.

Birnbaum RY, C. E., Agamy O, Kim MJ, Zhao J, Yamanaka T, Pappalardo Z, Clarke SL, Wenger AM, Nguyen L, Gurrieri F, Everman DB, Schwartz CE, Birk OS, Bejerano G, Lomvardas S, Ahituv N. (2012) 'Coding exons function as tissue-specific enhancers of nearby genes', *Genome Research* 22(6): 1059-68.

Boissan, M., Dabernat, S., Peuchant, E., Schlattner, U., Lascu, I. and Lacombe, M. L. (2009) 'The mammalian Nm23/NDPK family: from metastasis control to cilia movement', *Mol Cell Biochem* 329(1-2): 51-62.

Brewer, S., Feng, W., Huang, J., Sullivan, S. and Williams, T. (2004) 'Wnt1-Cre-mediated deletion of AP-2alpha causes multiple neural crest-related defects', *Dev Biol* 267(1): 135-52.

Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Yagle, M. K. and Palmiter, R. D. (1985) 'Factors affecting the efficiency of introducing foreign DNA into mice by microinjecting eggs', *Proc Natl Acad Sci U S A* 82(13): 4438-42.

Caspary, T., Larkins, C. E. and Anderson, K. V. (2007) 'The graded response to Sonic Hedgehog depends on cilia architecture', *Dev Cell* 12(5): 767-78.

Caubit, X., Thangarajah, R., Theil, T., Wirth, J., Nothwang, H. G., Ruther, U. and Krauss, S. (1999) 'Mouse Dac, a novel nuclear factor with homology to Drosophila dachshund shows a dynamic expression in the neural crest, the eye, the neocortex, and the limb bud', *Dev Dyn* 214(1): 66-80.

Chan, H. M. and La Thangue, N. B. (2001) 'p300/CBP proteins: HATs for transcriptional bridges and scaffolds', *J Cell Sci* 114(Pt 13): 2363-73.

Chandler, R. L., Chandler, K. J., McFarland, K. A. and Mortlock, D. P. (2007) 'Bmp2 transcription in osteoblast progenitors is regulated by a distant 3' enhancer located 156.3 kilobases from the promoter', *Mol Cell Biol* 27(8): 2934-51.

Chapman, D. L., Garvey, N., Hancock, S., Alexiou, M., Agulnik, S. I., Gibson-Brown, J. J., Cebra-Thomas, J., Bollag, R. J., Silver, L. M. and Papaioannou, V. E. (1996) 'Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development', *Dev Dyn* 206(4): 379-90.

Chen, C. H., Cretekos, C. J., Rasweiler, J. J. t. and Behringer, R. R. (2005) 'Hoxd13 expression in the developing limbs of the short-tailed fruit bat, Carollia perspicillata', *Evol Dev* 7(2): 130-41.

Cooper, K. L., Hu, J. K., ten Berge, D., Fernandez-Teran, M., Ros, M. A. and Tabin, C. J. (2011) 'Initiation of proximal-distal patterning in the vertebrate limb by signals and growth', *Science* 332(6033): 1083-6.

Cotney, J., Leng, J., Oh, S., DeMare, L. E., Reilly, S. K., Gerstein, M. B. and Noonan, J. P. (2012) 'Chromatin state signatures associated with tissue-specific gene expression and enhancer activity in the embryonic limb', *Genome Res* 22(6): 1069-80.

Cotney, J., Leng, J., Yin, J., Reilly, S. K., Demare, L. E., Emera, D., Ayoub, A. E., Rakic, P. and Noonan, J. P. (2013) 'The evolution of lineage-specific regulatory activities in the human embryonic limb', *Cell* 154(1): 185-96.

Cretekos, C. J., Wang, Y., Green, E. D., Martin, J. F., Rasweiler, J. J. t. and Behringer, R. R. (2008) 'Regulatory divergence modifies limb length between mammals', *Genes Dev* 22(2): 141-51.

Cretekos, C. J., Weatherbee, S. D., Chen, C. H., Badwaik, N. K., Niswander, L., Behringer, R. R. and Rasweiler, J. J. t. (2005) 'Embryonic staging system for the short-tailed fruit bat, Carollia perspicillata, a model organism for the mammalian order Chiroptera, based upon timed pregnancies in captive-bred animals', *Dev Dyn* 233(3): 721-38.

Dathe, K., Kjaer, K. W., Brehm, A., Meinecke, P., Nurnberg, P., Neto, J. C., Brunoni, D., Tommerup, N., Ott, C. E., Klopocki, E. et al. (2009) 'Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2', *Am J Hum Genet* 84(4): 483-92.

David, G., Grandinetti, K. B., Finnerty, P. M., Simpson, N., Chu, G. C. and Depinho, R. A. (2008) 'Specific requirement of the chromatin modifier mSin3B in cell cycle exit and cellular differentiation', *Proc Natl Acad Sci U S A* 105(11): 4168-72.

Davis, A. P., Witte, D. P., Hsieh-Li, H. M., Potter, S. S. and Capecchi, M. R. (1995) 'Absence of radius and ulna in mice lacking hoxa-11 and hoxd-11', *Nature* 375(6534): 791-5.

Davisson, M. T. and Cattanach, B. M. (1990) 'The mouse mutation ulnaless on chromosome 2', *J Hered* 81(2): 151-3.

Delaurier, A., Burton, N., Bennett, M., Baldock, R., Davidson, D., Mohun, T. J. and Logan, M. P. (2008) 'The Mouse Limb Anatomy Atlas: an interactive 3D tool for studying embryonic limb patterning', *BMC Dev Biol* 8: 83.

Donner, A. L. and Williams, T. (2006) 'Frontal nasal prominence expression driven by Tcfap2a relies on a conserved binding site for STAT proteins', *Dev Dyn* 235(5): 1358-70.

Echelard, Y., Epstein, D. J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. A. and McMahon, A. P. (1993) 'Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity', *Cell* 75(7): 1417-30.

Feng, W., Huang, J., Zhang, J. and Williams, T. (2008) 'Identification and analysis of a conserved Tcfap2a intronic enhancer element required for expression in facial and limb bud mesenchyme', *Mol Cell Biol* 28(1): 315-25.

Flynn, J. M., Meadows, E., Fiorotto, M. and Klein, W. H. (2010) 'Myogenin regulates exercise capacity and skeletal muscle metabolism in the adult mouse', *PLoS One* 5(10): e13535.

Fromental-Ramain, C., Warot, X., Messadecq, N., LeMeur, M., Dolle, P. and Chambon, P. (1996) 'Hoxa-13 and Hoxd-13 play a crucial role in the patterning of the limb autopod', *Development* 122(10): 2997-3011.

Furniss, D., Lettice, L. A., Taylor, I. B., Critchley, P. S., Giele, H., Hill, R. E. and Wilkie, A. O. (2008) 'A variant in the sonic hedgehog regulatory sequence (ZRS) is associated with triphalangeal thumb and deregulates expression in the developing limb', *Hum Mol Genet* 17(16): 2417-23.

Goetz, S. C., Ocbina, P. J. and Anderson, K. V. (2009) 'The primary cilium as a Hedgehog signal transduction machine', *Methods Cell Biol* 94: 199-222.

Gompel, N., Prud'homme, B., Wittkopp, P. J., Kassner, V. A. and Carroll, S. B. (2005) 'Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila', *Nature* 433(7025): 481-7.

Gonzalez, F., Duboule, D. and Spitz, F. (2007) 'Transgenic analysis of Hoxd gene regulation during digit development', *Dev Biol* 306(2): 847-59.

Guenther, C., Pantalena-Filho, L. and Kingsley, D. M. (2008) 'Shaping skeletal growth by modular regulatory elements in the Bmp5 gene', *PLoS Genet* 4(12): e1000308.

Haeussler, M. and Joly, J. S. (2011) 'When needles look like hay: how to find tissue-specific enhancers in model organism genomes', *Dev Biol* 350(2): 239-54.

Halligan, D. L., Oliver, F., Guthrie, J., Stemshorn, K. C., Harr, B. and Keightley, P. D. (2011) 'Positive and negative selection in murine ultraconserved noncoding elements', *Mol Biol Evol* 28(9): 2651-60.

Hamada, H., Meno, C., Watanabe, D. and Saijoh, Y. (2002) 'Establishment of vertebrate left-right asymmetry', *Nat Rev Genet* 3(2): 103-13.

Hasson, P., DeLaurier, A., Bennett, M., Grigorieva, E., Naiche, L. A., Papaioannou, V. E., Mohun, T. J. and Logan, M. P. 'Tbx4 and tbx5 acting in connective tissue are required for limb muscle and tendon patterning', *Dev Cell* 18(1): 148-56.

Haycraft, C. J., Banizs, B., Aydin-Son, Y., Zhang, Q., Michaud, E. J. and Yoder, B. K. (2005) 'Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function', *PLoS Genet* 1(4): e53.

Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A. et al. (2007) 'Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome', *Nat Genet* 39(3): 311-8.

Herault, Y., Fraudeau, N., Zakany, J. and Duboule, D. (1997) 'Ulnaless (Ul), a regulatory mutation inducing both loss-of-function and gain-of-function of posterior Hoxd genes', *Development* 124(18): 3493-500.

Herault, Y., Rassoulzadegan, M., Cuzin, F. and Duboule, D. (1998) 'Engineering chromosomes in mice through targeted meiotic recombination (TAMERE)', *Nat Genet* 20(4): 381-4.

Hill., R. E. (2007) 'How to make a zone of polarizing activity: insights into limb development via the abnormality preaxial polydactyly.', *Development, Growth, & Differentiation* 49(6): 439-448.

Hockman, D., Cretekos, C. J., Mason, M. K., Behringer, R. R., Jacobs, D. S. and Illing, N. (2008) 'A second wave of Sonic hedgehog expression during the development of the bat limb', *Proc Natl Acad Sci U S A* 105(44): 16982-7.

Hon, G. C., Hawkins, R. D. and Ren, B. (2009) 'Predictive chromatin signatures in the mammalian genome', *Hum Mol Genet* 18(R2): R195-201.

Husz, Z. L., Burton, N., Hill, B., Milyaev, N. and Baldock, R. A. (2012) 'Web tools for large-scale 3D biological images and atlases', *BMC Bioinformatics* 13: 122.

Kalkhoven, E. (2004) 'CBP and p300: HATs for different occasions', *Biochem Pharmacol* 68(6): 1145-55.

Ko, M. S. (2001) 'Embryogenomics: developmental biology meets genomics', *Trends Biotechnol* 19(12): 511-8.

Kolpakova-Hart, E., Jinnin, M., Hou, B., Fukai, N. and Olsen, B. R. (2007) 'Kinesin-2 controls development and patterning of the vertebrate skeleton by Hedgehog- and Gli3-dependent mechanisms', *Dev Biol* 309(2): 273-84.

Kong, Q., Zeng, W., Wu, J., Hu, W., Li, C. and Mao, B. (2010) 'RNF220, an E3 ubiquitin ligase that targets Sin3B for ubiquitination', *Biochem Biophys Res Commun* 393(4): 708-13.

Kothary, R., Clapoff, S., Darling, S., Perry, M. D., Moran, L. A. and Rossant, J. (1989) 'Inducible expression of an hsp68-lacZ hybrid gene in transgenic mice', *Development* 105(4): 707-14.

Kouwenhoven, E. N., van Heeringen, S. J., Tena, J. J., Oti, M., Dutilh, B. E., Alonso, M. E., de la Calle-Mustienes, E., Smeenk, L., Rinne, T., Parsaulian, L. et al. (2010) 'Genome-wide profiling of p63 DNA-binding sites identifies an element that regulates gene expression during limb development in the 7q21 SHFM1 locus', *PLoS Genet* 6(8): e1001065.

Laufer, E., Nelson, C. E., Johnson, R. L., Morgan, B. A. and Tabin, C. (1994) 'Sonic hedgehog and Fgf-4 act through a signaling cascade and feedback loop to integrate growth and patterning of the developing limb bud', *Cell* 79(6): 993-1003.

Lettice, L. A., Heaney, S. J., Purdie, L. A., Li, L., de Beer, P., Oostra, B. A., Goode, D., Elgar, G., Hill, R. E. and de Graaff, E. (2003) 'A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly', *Hum Mol Genet* 12(14): 1725-35.

Lettice, L. A., Horikoshi, T., Heaney, S. J., van Baren, M. J., van der Linde, H. C., Breedveld, G. J., Joosse, M., Akarsu, N., Oostra, B. A., Endo, N. et al. (2002) 'Disruption of a long-range cis-acting regulator for Shh causes preaxial polydactyly', *Proc Natl Acad Sci U S A* 99(11): 7548-53.

Levine, M. (2010) 'Transcriptional enhancers in animal development and evolution', *Curr Biol* 20(17): R754-63.

Li, Y. and Behringer, R. R. (1998) 'Esx1 is an X-chromosome-imprinted regulator of placental development and fetal growth', *Nat Genet* 20(3): 309-11.

Liu, Y. H., Ma, L., Wu, L. Y., Luo, W., Kundu, R., Sangiorgi, F., Snead, M. L. and Maxson, R. (1994) 'Regulation of the Msx2 homeobox gene during mouse embryogenesis: a transgene with 439 bp of 5' flanking sequence is expressed exclusively in the apical ectodermal ridge of the developing limb', *Mech Dev* 48(3): 187-97.

Lu, M. F., Cheng, H. T., Lacy, A. R., Kern, M. J., Argao, E. A., Potter, S. S., Olson, E. N. and Martin, J. F. (1999) 'Paired-related homeobox genes cooperate in handplate and hindlimb zeugopod morphogenesis', *Dev Biol* 205(1): 145-57.

Maas, R., Elfering, S., Glaser, T. and Jepeal, L. (1994) 'Deficient outgrowth of the ureteric bud underlies the renal agenesis phenotype in mice manifesting the limb deformity (ld) mutation', *Dev Dyn* 199(3): 214-28.

Maguire, C. T., Demarest, B. L., Hill, J. T., Palmer, J. D., Brothman, A. R., Yost, H. J. and Condic, M. L. (2013) 'Genome-wide analysis reveals the unique stem cell identity of human amniocytes', *PLoS One* 8(1): e53372.

Malik, S. and Roeder, R. G. (2010) 'The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation', *Nat Rev Genet* 11(11): 761-72.

Mansour, S. L., Thomas, K. R. and Capecchi, M. R. (1988) 'Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes', *Nature* 336(6197): 348-52.

Marinic, M., Aktas, T., Ruf, S. and Spitz, F. (2013) 'An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape', *Dev Cell* 24(5): 530-42.

Martin, J. F., Bradley, A. and Olson, E. N. (1995) 'The paired-like homeo box gene MHox is required for early events of skeletogenesis in multiple lineages', *Genes Dev* 9(10): 1237-49.

Martin, J. F. and Olson, E. N. (2000) 'Identification of a prx1 limb enhancer', *Genesis* 26(4): 225-9.

Masuya, H., Sagai, T., Wakana, S., Moriwaki, K. and Shiroishi, T. (1995) 'A duplicated zone of polarizing activity in polydactylous mouse mutants', *Genes Dev* 9(13): 1645-53.

Masuya, H., Sezutsu, H., Sakuraba, Y., Sagai, T., Hosoya, M., Kaneda, H., Miura, I., Kobayashi, K., Sumiyama, K., Shimizu, A. et al. (2007) 'A series of ENU-induced single-base substitutions in a long-range cis-element altering Sonic hedgehog expression in the developing mouse limb bud', *Genomics* 89(2): 207-14.

Meadows, E., Flynn, J. M. and Klein, W. H. (2011) 'Myogenin regulates exercise capacity but is dispensable for skeletal muscle regeneration in adult mdx mice', *PLoS One* 6(1): e16184.

Mendelsohn, C., Ruberte, E., LeMeur, M., Morriss-Kay, G. and Chambon, P. (1991) 'Developmental analysis of the retinoic acid-inducible RAR-beta 2 promoter in transgenic animals', *Development* 113(3): 723-34.

Menke, D. B., Guenther, C. and Kingsley, D. M. (2008) 'Dual hindlimb control elements in the Tbx4 gene and region-specific control of bone size in vertebrate limbs', *Development* 135(15): 2543-53.

Mitgutsch, C., Richardson, M. K., Jimenez, R., Martin, J. E., Kondrashov, P., de Bakker, M. A. and Sanchez-Villagra, M. R. (2012) 'Circumventing the polydactyly 'constraint': the mole's 'thumb'', *Biol Lett* 8(1): 74-7.

Mohan, K. N., Ding, F. and Chaillet, J. R. (2011) 'Distinct roles of DMAP1 in mouse development', *Mol Cell Biol* 31(9): 1861-9.

Montavon, T., Soshnikova, N., Mascrez, B., Joye, E., Thevenet, L., Splinter, E., de Laat, W., Spitz, F. and Duboule, D. (2011) 'A regulatory archipelago controls Hox genes transcription in digits', *Cell* 147(5): 1132-45.

Mortlock, D. P., Guenther, C. and Kingsley, D. M. (2003) 'A general approach for identifying distant regulatory elements applied to the Gdf6 gene', *Genome Res* 13(9): 2069-81.

Murthi, P., Kalionis, B., Rajaraman, G., Keogh, R. J. and Da Silva Costa, F. (2012) 'The role of homeobox genes in the development of placental insufficiency', *Fetal Diagn Ther* 32(4): 225-30.

Nagy, A., Gertsenstein, M., Vintersten, K. and Behringer, R. (2003) 'Manipulating the Mouse Embryo: A Laboratory Manual', *Cold Springs Harbor Laboratory Press*.

Naiche, L. A., Arora, R., Kania, A., Lewandoski, M. and Papaioannou, V. E. 'Identity and fate of Tbx4-expressing cells reveal developmental cell fate decisions in the allantois, limb, and external genitalia', *Dev Dyn* 240(10): 2290-300.

Naiche, L. A. and Papaioannou, V. E. (2003) 'Loss of Tbx4 blocks hindlimb development and affects vascularization and fusion of the allantois', *Development* 130(12): 2681-93.

Naiche, L. A. and Papaioannou, V. E. (2007) 'Tbx4 is not required for hindlimb identity or post-bud hindlimb outgrowth', *Development* 134(1): 93-103.

Niswander, L., Tickle, C., Vogel, A., Booth, I. and Martin, G. R. (1993) 'FGF-4 replaces the apical ectodermal ridge and directs outgrowth and patterning of the limb', *Cell* 75(3): 579-87.

Nobrega, M. A., Ovcharenko, I., Afzal, V. and Rubin, E. M. (2003) 'Scanning human gene deserts for long-range enhancers', *Science* 302(5644): 413.

Nolte, M. J., Hockman, D., Cretekos, C. J., Behringer, R. R. and Rasweiler, J. J. t. (2009) 'Embryonic staging system for the Black Mastiff Bat, Molossus rufus (Molossidae), correlated with structure-function relationships in the adult', *Anat Rec (Hoboken)* 292(2): 155-68, spc 1.

Nottoli, T., Hagopian-Donaldson, S., Zhang, J., Perkins, A. and Williams, T. (1998) 'AP-2-null cells disrupt morphogenesis of the eye, face, and limbs in chimeric mice', *Proc Natl Acad Sci U S A* 95(23): 13714-9.

O'Gorman, S., Dagenais, N. A., Qian, M. and Marchuk, Y. (1997) 'Protamine-Cre recombinase transgenes efficiently recombine target sequences in the male germ line of mice, but not in embryonic stem cells', *Proc Natl Acad Sci U S A* 94(26): 14602-7.

Ohler, U. and Wassarman, D. A. (2010) 'Promoting developmental transcription', *Development* 137(1): 15-26.

Ong, C. T. and Corces, V. G. (2011) 'Enhancer function: new insights into the regulation of tissue-specific gene expression', *Nat Rev Genet* 12(4): 283-93.

Ovcharenko, I., Nobrega, M. A., Loots, G. G. and Stubbs, L. (2004) 'ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes', *Nucleic Acids Res* 32(Web Server issue): W280-6.

Pajni-Underwood, S., Wilson, C. P., Elder, C., Mishina, Y. and Lewandoski, M. (2007) 'BMP signals control limb bud interdigital programmed cell death by regulating FGF signaling', *Development* 134(12): 2359-68.

Peichel, C. L., Prabhakaran, B. and Vogt, T. F. (1997) 'The mouse Ulnaless mutation deregulates posterior HoxD gene expression and alters appendicular patterning', *Development* 124(18): 3481-92.

Pennacchio, L. A., Ahituv, N., Moses, A. M., Prabhakar, S., Nobrega, M. A., Shoukry, M., Minovitsky, S., Dubchak, I., Holt, A., Lewis, K. D. et al. (2006) 'In vivo enhancer analysis of human conserved non-coding sequences', *Nature* 444(7118): 499-502.

Pennacchio, L. A. and Rubin, E. M. (2001) 'Genomic strategies to identify mammalian regulatory sequences', *Nat Rev Genet* 2(2): 100-9.

Prabhakar, S., Poulin, F., Shoukry, M., Afzal, V., Rubin, E. M., Couronne, O. and Pennacchio, L. A. (2006) 'Close sequence comparisons are sufficient to identify human cis-regulatory elements', *Genome Res* 16(7): 855-63.

Prabhakar, S., Visel, A., Akiyama, J. A., Shoukry, M., Lewis, K. D., Holt, A., Plajzer-Frick, I., Morrison, H., Fitzpatrick, D. R., Afzal, V. et al. (2008) 'Human-specific gain of function in a developmental enhancer', *Science* 321(5894): 1346-50.

Rabinowitz, A. H. and Vokes, S. A. (2012) 'Integration of the transcriptional networks regulating limb morphogenesis', *Dev Biol* 368(2): 165-80.

Ramirez-Solis, R., Rivera-Perez, J., Wallace, J. D., Wims, M., Zheng, H. and Bradley, A. (1992) 'Genomic DNA microextraction: a method to screen numerous samples', *Anal Biochem* 201(2): 331-5.

Ray, R. and Capecchi, M. (2008) 'An examination of the Chiropteran HoxD locus from an evolutionary perspective', *Evol Dev* 10(6): 657-70.

Reynolds, K., Mezey, E. and Zimmer, A. (1991) 'Activity of the beta-retinoic acid receptor promoter in transgenic mice', *Mech Dev* 36(1-2): 15-29.

Riddle, R. D., Johnson, R. L., Laufer, E. and Tabin, C. (1993) 'Sonic hedgehog mediates the polarizing activity of the ZPA', *Cell* 75(7): 1401-16.

Rodriguez, T. A., Sparrow, D. B., Scott, A. N., Withington, S. L., Preis, J. I., Michalicek, J., Clements, M., Tsang, T. E., Shioda, T., Beddington, R. S. et al. (2004) 'Cited1 is required in trophoblasts for placental development and for embryo growth and survival', *Mol Cell Biol* 24(1): 228-44.

Rosello-Diez, A., Ros, M. A. and Torres, M. (2011) 'Diffusible signals, not autonomous mechanisms, determine the main proximodistal limb subdivision', *Science* 332(6033): 1086-8.

Ruf, S., Symmons, O., Uslu, V. V., Dolle, D., Hot, C., Ettwiller, L. and Spitz, F. 'Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor', *Nat Genet* 43(4): 379-86.

Sagai, T., Hosoya, M., Mizushina, Y., Tamura, M. and Shiroishi, T. (2005) 'Elimination of a long-range cis-regulatory module causes complete loss of limb-specific Shh expression and truncation of the mouse limb', *Development* 132(4): 797-803.

Sagai, T., Masuya, H., Tamura, M., Shimizu, K., Yada, Y., Wakana, S., Gondo, Y., Noda, T. and Shiroishi, T. (2004) 'Phylogenetic conservation of a limb-specific, cis-acting regulator of Sonic hedgehog (Shh)', *Mamm Genome* 15(1): 23-34.

Sakabe, N. J., Savic, D. and Nobrega, M. A. (2012) 'Transcriptional enhancers in development and disease', *Genome Biol* 13(1): 238.

Sanchez-Villagra, M. R. and Menke, P. R. (2005) 'The mole's thumb -- evolution of the hand skeleton in talpids (Mammalia)', *Zoology (Jena)* 108(1): 3-12.

Sanyal, A., Lajoie, B. R., Jain, G. and Dekker, J. (2012) 'The long-range interaction landscape of gene promoters', *Nature* 489(7414): 109-13.

Schorle, H., Meier, P., Buchert, M., Jaenisch, R. and Mitchell, P. J. (1996) 'Transcription factor AP-2 essential for cranial closure and craniofacial development', *Nature* 381(6579): 235-8.

Schughart, K., Bieberich, C. J., Eid, R. and Ruddle, F. H. (1991) 'A regulatory region from the mouse Hox-2.2 promoter directs gene expression into developing limbs', *Development* 112(3): 807-11.

Shan, S. W., Tang, M. K., Cai, D. Q., Chui, Y. L., Chow, P. H., Grotewold, L. and Lee, K. K. (2005) 'Comparative proteomic analysis identifies protein disulfide isomerase and peroxiredoxin 1 as new players involved in embryonic interdigital cell death', *Dev Dyn* 233(2): 266-81.

Sharpe, J., Lettice, L., Hecksher-Sorensen, J., Fox, M., Hill, R. and Krumlauf, R. (1999) 'Identification of sonic hedgehog as a candidate gene responsible for the polydactylous mouse mutant Sasquatch', *Curr Biol* 9(2): 97-100.

Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V. V. et al. (2012) 'A map of the cis-regulatory sequences in the mouse genome', *Nature* 488(7409): 116-20.

Shubin, N., Tabin, C. and Carroll, S. (1997) 'Fossils, genes and the evolution of animal limbs', *Nature* 388(6643): 639-48.

Shubin, N., Tabin, C. and Carroll, S. (2009) 'Deep homology and the origins of evolutionary novelty', *Nature* 457(7231): 818-23.

Shubin, N. H., Daeschler, E. B. and Jenkins, F. A., Jr. (2006) 'The pectoral fin of Tiktaalik roseae and the origin of the tetrapod limb', *Nature* 440(7085): 764-71.

Spitz, F. and Furlong, E. E. (2012) 'Transcription factors: from enhancer binding to developmental control', *Nat Rev Genet* 13(9): 613-26.

Spitz, F., Gonzalez, F. and Duboule, D. (2003) 'A global control region defines a chromosomal regulatory landscape containing the HoxD cluster', *Cell* 113(3): 405-17.

Spitz, F., Gonzalez, F., Peichel, C., Vogt, T. F., Duboule, D. and Zakany, J. (2001) 'Large scale transgenic and cluster deletion analysis of the HoxD complex separate an ancestral regulatory module from evolutionary innovations', *Genes Dev* 15(17): 2209-14.

Spitz, F., Herkenne, C., Morris, M. A. and Duboule, D. (2005) 'Inversion-induced disruption of the Hoxd cluster leads to the partition of regulatory landscapes', *Nat Genet* 37(8): 889-93.

Sun, M., Ma, F., Zeng, X., Liu, Q., Zhao, X. L., Wu, F. X., Wu, G. P., Zhang, Z. F., Gu, B., Zhao, Y. F. et al. (2008) 'Triphalangeal thumb-polysyndactyly syndrome and syndactyly type IV are caused by genomic duplications involving the long range, limb-specific SHH enhancer', *J Med Genet* 45(9): 589-95.

Sun, X., Mariani, F. V. and Martin, G. R. (2002) 'Functions of FGF signalling from the apical ectodermal ridge in limb development', *Nature* 418(6897): 501-8.

Sundararajan, S., Wakamiya, M., Behringer, R. R. and Rivera-Perez, J. A. (2012) 'A fast and sensitive alternative for beta-galactosidase detection in mouse embryos', *Development* 139(23): 4484-90.

Taher, L., Collette, N. M., Murugesh, D., Maxwell, E., Ovcharenko, I. and Loots, G. G. (2011a) 'Global gene expression analysis of murine limb development', *PLoS One* 6(12): e28358.

Taher, L., McGaughey, D. M., Maragh, S., Aneas, I., Bessling, S. L., Miller, W., Nobrega, M. A., McCallion, A. S. and Ovcharenko, I. (2011b) 'Genome-wide identification of conserved regulatory function in diverged sequences', *Genome Res* 21(7): 1139-49.

Tarchini, B. and Duboule, D. (2006) 'Control of Hoxd genes' collinearity during early limb development', *Dev Cell* 10(1): 93-103.

Tickle, C. (2006) 'Making digit patterns in the vertebrate limb', *Nat Rev Mol Cell Biol* 7(1): 45-53.

VanderMeer, J. E. and Ahituv, N. 'cis-regulatory mutations are a genetic cause of human limb malformations', *Dev Dyn* 240(5): 920-30.

Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F. et al. (2009a) 'ChIP-seq accurately predicts tissue-specific activity of enhancers', *Nature* 457(7231): 854-8.

Visel, A., Bristow, J. and Pennacchio, L. A. (2007a) 'Enhancer identification through comparative genomics', *Semin Cell Dev Biol* 18(1): 140-52.

Visel, A., Minovitsky, S., Dubchak, I. and Pennacchio, L. A. (2007b) 'VISTA Enhancer Browser--a database of tissue-specific human enhancers', *Nucleic Acids Res* 35(Database issue): D88-92.

Visel, A., Prabhakar, S., Akiyama, J. A., Shoukry, M., Lewis, K. D., Holt, A., Plajzer-Frick, I., Afzal, V., Rubin, E. M. and Pennacchio, L. A. (2008) 'Ultraconservation identifies a small subset of extremely constrained developmental enhancers', *Nat Genet* 40(2): 158-60.

Visel, A., Rubin, E. M. and Pennacchio, L. A. (2009b) 'Genomic views of distant-acting enhancers', *Nature* 461(7261): 199-205.

Vogel, P., Read, R., Hansen, G. M., Freay, L. C., Zambrowicz, B. P. and Sands, A. T. (2010) 'Situs inversus in Dpcd/Poll-/-, Nme7-/- , and Pkd1l1-/- mice', *Vet Pathol* 47(1): 120-31.

Vokes, S. A., Ji, H., Wong, W. H. and McMahon, A. P. (2008) 'A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb', *Genes Dev* 22(19): 2651-63.

Wang, C. C., Chan, D. C. and Leder, P. (1997) 'The mouse formin (Fmn) gene: genomic structure, novel exons, and genetic mapping', *Genomics* 39(3): 303-11.

Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I. and Young, R. A. (2013) 'Master transcription factors and mediator establish super-enhancers at key cell identity genes', *Cell* 153(2): 307-19.

Wieczorek, D., Pawlik, B., Li, Y., Akarsu, N. A., Caliebe, A., May, K. J., Schweiger, B., Vargas, F. R., Balci, S., Gillessen-Kaesbach, G. et al. 'A specific mutation in the distant sonic hedgehog (SHH) cis-regulator (ZRS) causes Werner mesomelic syndrome (WMS) while complete ZRS duplications underlie Haas type polysyndactyly and preaxial polydactyly (PPD) with or without triphalangeal thumb', *Hum Mutat* 31(1): 81-9.

Woltering, J. M. and Duboule, D. (2010) 'The origin of digits: expression patterns versus regulatory mechanisms', *Dev Cell* 18(4): 526-32.

Wong, S. Y., Seol, A. D., So, P. L., Ermilov, A. N., Bichakjian, C. K., Epstein, E. H., Jr., Dlugosz, A. A. and Reiter, J. F. (2009) 'Primary cilia can both mediate and suppress Hedgehog pathway-dependent tumorigenesis', *Nat Med* 15(9): 1055-61.

Xi, H., Shulha, H. P., Lin, J. M., Vales, T. R., Fu, Y., Bodine, D. M., McKay, R. D., Chenoweth, J. G., Tesar, P. J., Furey, T. S. et al. (2007) 'Identification and characterization of

cell type-specific and ubiquitous chromatin regulatory structures in the human genome', *PLoS Genet* 3(8): e136.

Zakany, J. and Duboule, D. (1996) 'Synpolydactyly in mice with a targeted deficiency in the HoxD complex', *Nature* 384(6604): 69-71.

Zakany, J. and Duboule, D. (2007) 'The role of Hox genes during vertebrate limb development', *Curr Opin Genet Dev* 17(4): 359-66.

Zakany, J., Fromental-Ramain, C., Warot, X. and Duboule, D. (1997) 'Regulation of number and size of digits by posterior Hox genes: a dose-dependent mechanism with potential evolutionary implications', *Proc Natl Acad Sci U S A* 94(25): 13695-700.

Zakany, J., Kmita, M. and Duboule, D. (2004) 'A dual role for Hox genes in limb anterior-posterior asymmetry', *Science* 304(5677): 1669-72.

Zeller, R., Haramis, A. G., Zuniga, A., McGuigan, C., Dono, R., Davidson, G., Chabanis, S. and Gibson, T. (1999) 'Formin defines a large family of morphoregulatory genes and functions in establishment of the polarising region', *Cell Tissue Res* 296(1): 85-93.

Zeller, R., Lopez-Rios, J. and Zuniga, A. (2009) 'Vertebrate limb bud development: moving towards integrative analysis of organogenesis', *Nat Rev Genet* 10(12): 845-58.

Zhang, J., Hagopian-Donaldson, S., Serbedzija, G., Elsemore, J., Plehn-Dujowich, D., McMahon, A. P., Flavell, R. A. and Williams, T. (1996) 'Neural tube, skeletal and body wall defects in mice lacking transcription factor AP-2', *Nature* 381(6579): 238-41.

Zhang, J. and Williams, T. (2003) 'Identification and regulation of tissue-specific cis-acting elements associated with the human AP-2alpha gene', *Dev Dyn* 228(2): 194-207.

Zuniga, A., Michos, O., Spitz, F., Haramis, A. P., Panman, L., Galli, A., Vintersten, K., Klasen, C., Mansfield, W., Kuc, S. et al. (2004) 'Mouse limb deformity mutations disrupt a global control region within the large regulatory landscape required for Gremlin expression', *Genes Dev* 18(13): 1553-64.

VITA

Mark Joseph Nolte was born in southern California on May 11, 1980, the fourth son of Joyce Ann Nothum Nolte and Gregory Robert Nolte. After graduating from Poway High School, Poway, California, in 1998 he entered Brigham Young University in Provo, Utah. In December 2005 he received two Bachelors of Science with majors in Molecular Biology and Geology, and a minor in Communications (Print Journalism). From 1999-2001 he served a mission for The Church of Jesus Christ of Latter-day Saints in Fortaleza, Brazil. In 2006 he entered The University of Texas Health Science Center at Houston Graduate School of Biomedical Sciences. He is happily married to Emily Taye Nolte and they are the grateful parents of four children, two of which were born in Houston.

Permanent address:

136 S. 400 East

Springville, UT 84663